B14

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number
WO 03/014303 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/24655

(22) International Filing Date: 2 August 2002 (02.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/309,841 3 August 2001 (03.08.2001) US
60/360,061 25 February 2002 (25.02.2002) US

(71) Applicant *(for all designated States except US)*: **ARBOR VITA CORPORATION** [US/US]; 772 Lucerne Drive, Sunnyvale, CA 94085 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **LU, Peter, S.** [US/US]; 99 East Middlefield Road, No. 29, Mountain View, CA 94043 (US). **RABINOWITZ, Joshua, D.** [US/US]; 750 N. Shoreline Blvd., No. 50, Mountain View, CA 94043 (US). **SCHWEIZER, Johannes** [DE/US]; 284 Tyrella Avenue, No. 17, Mountain View, CA 94043 (US). **CARRICK, Deanna, Marie** [US/US]; 3584 Lancelot Court, Fremont, CA 94536 (US).

(74) Agents: **AUSENHUS, Scott, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

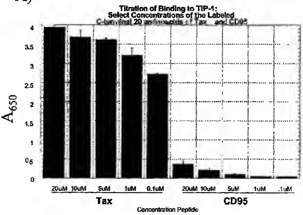
(81) Designated States *(national)*: AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

[Continued on next page]

(54) Title: MOLECULAR INTERACTIONS IN CELLS

A)



(57) Abstract: The invention provides reagents and methods for inhibiting or enhancing interactions between proteins in cells, particularly interactions between a PDZ protein and a PL protein. Reagents and methods that are provided are useful for treatment of a variety of diseases and conditions in a variety of cell types.

WO 03/014303 A2

European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MOLECULAR INTERACTIONS IN CELLS

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application claims the benefit of U.S. Provisional Application No. 60/309,841, filed August 3, 2001 and U.S. Provisional Application No. 60/360,061, filed February 25, 2002, each of which is incorporated herein by reference in its entirety for all purposes.

FIELD OF THE INVENTION

Peptides and peptide analogues, and methods for using such compositions, to regulate various biological functions of cells are provided. For example, certain peptides and peptide analogues which are provided are utilized in methods for modulating a biological function in certain cells by antagonizing or promoting binding between a protein having a PDZ domain and a protein that binds a PDZ domain. Also provided are methods for identifying compounds that modulate the interactions between specific PDZ domains and their ligands.

BACKGROUND

PDZ domains of proteins are named after three prototypical proteins: post-synaptic density protein 95 (PSD95), Drosophila large disc protein (Dlg1) and Zonula Occludin 1 protein (ZO-1; Gomperts et al., 1996, *Cell* 84:659-662). PDZ domains contain the signature sequence GLGF. The first PDZ proteins were identified as functioning to concentrate receptors at neuronal synapses or tight junctions. In the nervous system, typical PDZ domain-containing proteins contain three PDZ domains, one SH3 domain and one guanylate kinase domain. Examples of intracellular PDZ domain-containing proteins include LIN-2, LIN-7 and LIN-10 at the pre-synapse, and PSD95 at the post-synapse. PDZ domains have been shown to bind the carboxyl termini of transmembrane proteins in neuronal cells. Songyang et al. reported that proteins capable of binding PDZ domains contain a carboxyl terminal motif sequence of E-S/T-X-V/I (Songyang et al., 1997, *Science* 275:73). X-ray crystallography studies have revealed the contact points between the motif sequence and PDZ domains (Doyle et al., 1996, *Cell* 88:1067-1076).

The role of PDZ domain:PDZ ligand (PL) interactions in human disease has only recently begun to be studied. Deletions that remove the PL of the human Cystic Fibrosis

Transmembrane Conductance regulator (CFTR) have been correlated with an increase in Cystic Fibrosis and underscore the importance of proper PDZ:PL function (Benharouga et al 2001, J. Cell. Biol. 153:957-70). Mouse gene disruptions in the PDZ domain-containing protein Shroom result in neural tube defects, a precursor to such disorders as exencephaly, acrania, facial clefting and spina bifida (Hildebrand and Soriano, 1999, Cell 99:485-497). In a similar manner, knockout mice at the Cypher gene locus (another PDZ domain-containing protein) result in a severe form of congenital myopathy and post-natally (Zhou et al 2001, J. Cell Biol. 155:605-12).

Given the paucity of information regarding the role that PDZ proteins play in biological functions and their role in disease, further information on interactions involving proteins with PDZ domains would be useful in understanding a number of different biological functions in cells and for the treatment of human disorders.

SUMMARY

Methods and compositions for modulating biological function in a variety of cell types (e.g., hematopoietic, neuronal, brain, stem, epidermal and epithelial) are provided herein. These methods and compositions can be utilized to treat various maladies including, but not limited to, diseases such as immune disorders, nervous system disorders and muscle disorders, for example. More specifically, these methods and compositions are for modulating binding between certain PDZ proteins and PL protein binding pairs as shown in **TABLE 7**. Other methods and compositions are for modulating binding between PDZ protein and PL protein binding pairs as listed in **TABLE 12**.

Certain methods involve introducing into the cell an agent that alters binding between a PDZ protein and a PL protein in the cell, whereby the biological function is modulated in the cell, and wherein the PDZ protein and PL protein are a binding pair as specified in **TABLE 7** or **TABLE 12**. In some of these methods, the agent is a polypeptide comprising at least the two, three or four carboxy-terminal residues of the PL protein.

The PDZ proteins and PL proteins that have been identified as interacting can be classified into a number of different groups, and provide an indication of the diverse functions that can be modulated using the methods and compounds that are provided herein. For example, the PDZ proteins can be: 1) an enzyme such as a protein kinase, a guanalyte kinase, a tyrosine phosphatase or a serine phosphatase, 2) a LIM protein, 3) a guanine

exchange factor, or 4) a viral oncogene interacting protein. Likewise, PL proteins can be 1) a T-cell surface receptor or a B-cell surface receptor, 2) a natural killer surface receptor, a monocyte cell surface receptor, or a granulocyte cell surface receptor, 3) an endothelial cell surface receptor, 4) a G-protein linked receptor or a regulator of G-protein signaling, 5) an adhesion protein or a tight junction integral membrane protein, 6) a viral oncogene, 7) neuron membrane transport protein, 8) a receptor kinase, 9) an ion channel or transporter protein, or 10) a tumor suppressor protein.

Modulation can be conducted in vitro or in vivo. If done in vitro, the cell into which the agent is introduced can be a cell within a cell culture.

Screening methods to identify compounds that modulate binding between PDZ proteins and PL peptides or proteins are also provided. Some screening methods involve contacting under suitable binding conditions (i) a PDZ–domain polypeptide having a sequence from a PDZ protein, and (ii) a PL peptide, wherein the PL peptide comprises a C-terminal sequence of the PL protein, the PDZ –domain polypeptide and the PL peptide are a binding pair as specified in **TABLES 7 or 12**; and contacting is performed in the presence of the test compound. Presence or absence of complex is then detected. The presence of the complex at a level that is statistically significantly higher in the presence of the test compound than in the absence of test compound is an indication that the test compound is an agonist, whereas, the presence of the complex at a level that is statistically significantly lower in the presence of the test compound than in the absence of test compound is an indication that the test compound is an antagonist.

Modulators of binding between a PDZ protein and a PL protein are also described herein. In certain instances, the modulator is (a) a peptide comprising at least 3 residues of a C-terminal sequence of a PL protein, and wherein the PDZ protein and the PL protein are a binding pair as specified in **TABLES 7 or 12**; or (b) a peptide mimetic of the peptide of section (a); or (c) a small molecule having similar functional activity with respect to the PDZ and PL protein binding pair as the peptide of section (a). The modulator can be either an agonist or antagonist. Such modulators can be formulated as a pharmaceutical composition.

Methods of treating a disease correlated with binding between a PDZ protein and a PL protein are also disclosed herein, the method comprising administering a therapeutically effective amount of a modulator as provided herein, wherein the PDZ

protein and the PL protein are a binding pair as specified in **TABLES 7** or **12**. As indicated supra, such methods can be used to treat a variety of diseases including, but not limited to, neurological disease, an immune response disease, a muscular disease, or a cancer. The methods can be used to treat humans and non-human animals, including for example, cattle, swine, sheep, dogs, cats, horses and the like.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURES 1A and 1B shows the results of introduction of a Tat-CD3 fusion peptide on T cell activation. Antigen-specific T cell activation was measured by cytokine production. Fusion peptides containing tat and a T cell surface molecule carboxyl terminus inhibited γ-interferon (IFN) production by a T cell line in response to myelin basic protein (MBP) stimulation. The level of inhibition was determined by first subtracting the binding of the labeled peptide to GST alone from the binding to the fusion protein and dividing by the signal in the absence of competitor peptide.

FIGURES 2A, 2B and **2C** show binding and competition assays with the PDZ ligands of CD95 (Fas) and Tax for the PDZ domain of TIP-1. **FIGURE 2A** shows a titration of Tax and CD95 PDZ ligands against a constant amount of TIP-1 protein. **FIGURE 2B** shows the ability of an unlabeled 8 amino acid peptide corresponding to the C-terminus of Tax to inhibit the binding of 20uM CD95 to TIP-1. **FIGURE 2C** shows the ability of an unlabeled 8 amino acid peptide corresponding to the C-terminus of CD95 to inhibit the binding of 1 uM Tax to TIP-1.

DESCRIPTION

I. Definitions

A "fusion protein" or "fusion polypeptide" as used herein refers to a composite protein, i.e., a single contiguous amino acid sequence, made up of two (or more) distinct, heterologous polypeptides that are not normally fused together in a single amino acid sequence. Thus, a fusion protein can include a single amino acid sequence that contains two entirely distinct amino acid sequences or two similar or identical polypeptide sequences, provided that these sequences are not normally found together in the same configuration in a single amino acid sequence found in nature. Fusion proteins can generally be prepared using either

recombinant nucleic acid methods, i.e., as a result of transcription and translation of a recombinant gene fusion product, which fusion comprises a segment encoding a polypeptide of the invention and a segment encoding a heterologous protein, or by chemical synthesis methods well known in the art.

A "fusion protein construct" as used herein is a polynucleotide encoding a fusion protein.

As used herein, the term "PDZ domain" refers to protein sequence (i.e., modular protein domain) of approximately 90 amino acids, characterized by homology to the brain synaptic protein PSD-95, the Drosophila septate junction protein Discs-Large (DLG), and the epithelial tight junction protein ZO1 (ZO1). PDZ domains are also known as Discs-Large homology repeats ("DHRs") and GLGF repeats. PDZ domains generally appear to maintain a core consensus sequence (Doyle, D. A., 1996, *Cell* 85: 1067-76).

PDZ domains are found in diverse membrane-associated proteins including members of the MAGUK family of guanylate kinase homologs, several protein phosphatases and kinases, neuronal nitric oxide synthase, and several dystrophin-associated proteins, collectively known as syntrophins.

Exemplary PDZ domain-containing proteins and PDZ domain sequences are shown in **TABLE 9**. The term "PDZ domain" also encompasses variants (e.g., naturally occurring variants) of the sequences of **TABLE 9** (e.g., polymorphic variants, variants with conservative substitutions, and the like). Typically, PDZ domains are substantially identical to those shown in **TABLE 9**, e.g., at least about 70%, at least about 80%, or at least about 90% amino acid residue identity when compared and aligned for maximum correspondence.

As used herein, the term "PDZ protein" refers to a naturally occurring protein containing a PDZ domain. Exemplary PDZ proteins include CASK, MPP1, DLG1, PSD95, NeDLG, TIP-33, SYN1a, TIP-43, LDP, LIM, LIMK1, LIMK2, MPP2, NOS1, AF6, PTN-4, prIL16, 41.8kD, KIAA0559, RGS12, KIAA0316, DVL1, TIP-40, TIAM1, MINT1, KIAA0303, CBP, MINT3, TIP-2, KIAA0561, and those listed in **TABLE 9**.

As used herein, the term "PDZ-domain polypeptide" refers to a polypeptide containing a PDZ domain, such as a fusion protein including a PDZ domain sequence, a naturally occurring PDZ protein, or an isolated PDZ domain peptide.

As used herein, the term "PL protein" or "PDZ Ligand protein" refers to a naturally occurring protein that forms a molecular complex with a PDZ-domain, or to a protein whose carboxy-terminus, when expressed separately from the full length protein (e.g., as a peptide fragment of 4-25 residues, e.g., 8, 10, 12, 14 or 16 residues), forms such a molecular complex. The molecular complex can be observed *in vitro* using the "A assay" or "G assay" described *infra*, or *in vivo*. Exemplary PL proteins listed in **TABLE 8** are demonstrated to bind specific PDZ proteins. This definition is not intended to include anti-PDZ antibodies and the like.

As used herein, a "PL sequence" refers to the amino acid sequence of the C-terminus of a PL protein (e.g., the C-terminal 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 20 or 25 residues) ("C-terminal PL sequence") or to an internal sequence known to bind a PDZ domain ("internal PL sequence").

As used herein, a "PL peptide" is a peptide of having a sequence from, or based on, the sequence of the C-terminus of a PL protein. Exemplary PL peptides (biotinylated) are listed in **TABLE 8**.

As used herein, a "PL fusion protein" is a fusion protein that has a PL sequence as one domain, typically as the C-terminal domain of the fusion protein. An exemplary PL fusion protein is a tat-PL sequence fusion.

As used herein, the term "PL inhibitor peptide sequence" refers to a PL peptide amino acid sequence that (in the form of a peptide or PL fusion protein) inhibits the interaction between a PDZ domain polypeptide and a PL peptide (e.g., in an A assay or a G assay).

As used herein, a "PDZ-domain encoding sequence" means a segment of a polynucleotide encoding a PDZ domain. In various embodiments, the polynucleotide is DNA, RNA, single stranded or double stranded.

As used herein, the terms "antagonist" and "inhibitor," when used in the context of modulating a binding interaction (such as the binding of a PDZ domain sequence to a PL sequence), are used interchangeably and refer to an agent that reduces the binding of the, e.g., PL sequence (e.g., PL peptide) and the, e.g., PDZ domain sequence (e.g., PDZ protein, PDZ domain peptide).

As used herein, the terms "agonist" and "enhancer," when used in the context of modulating a binding interaction (such as the binding of a PDZ domain sequence to a PL sequence), are used interchangeably and refer to an agent that increases the binding of the, e.g., PL sequence (e.g., PL peptide) and the, e.g., PDZ domain sequence (e.g., PDZ protein, PDZ domain peptide).

As used herein, the terms "peptide mimetic, " "peptidomimetic," and "peptide analog" are used interchangeably and refer to a synthetic chemical compound that has substantially the same structural and/or functional characteristics of a PL inhibitory or PL binding peptide of the invention. The mimetic can be either entirely composed of synthetic, non-natural analogues of amino acids, or, is a chimeric molecule of partly natural peptide amino acids and partly non-natural analogs of amino acids. The mimetic can also incorporate any amount of natural amino acid conservative substitutions as long as such substitutions also do not substantially alter the mimetic's structure and/or inhibitory or binding activity. As with polypeptides of the invention which are conservative variants, routine experimentation will determine whether a mimetic is within the scope of the invention, i.e., that its structure and/or function is not substantially altered. Thus, a mimetic composition is within the scope of the invention if it is capable of binding to a PDZ domain and/or inhibiting a PL-PDZ interaction.

Polypeptide mimetic compositions can contain any combination of nonnatural structural components, which are typically from three structural groups: a) residue linkage groups other than the natural amide bond ("peptide bond") linkages; b) non-natural residues in place of naturally occurring amino acid residues; or c) residues which induce secondary

structural mimicry, i.e., to induce or stabilize a secondary structure, e.g., a beta turn, gamma turn, beta sheet, alpha helix conformation, and the like.

A polypeptide can he characterized as a mimetic when all or some of its residues are joined hy chemical means other than natural peptide honds. Individual peptidomimetic residues can be joined by peptide honds, other chemical bonds or coupling means, such as, e.g., glutaraldehyde, N-hydroxysuccinimide esters, bifunctional maleimides, N,N=-dicyclohexylcarbodiimide (DCC) or N,N=-diisopropylcarhodiimide (DIC). Linking groups that can he an alternative to the traditional amide bond ("peptide hond") linkages include, e.g., ketomethylene (e.g., -C(=O)-CH_2- for -C(=O)-NH-), aminomethylene (CH_2-NH), ethylene, olefin (CH=CH), ether (CH_2-O), thioether (CH_2-S), tetrazole (CN_4-), thiazole, retroamide, thioamide, or ester (see, e.g., Spatola (1983) in Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, A Peptide Backbone Modifications, Marcell Delkker, NY).

A polypeptide can also be characterized as a mimetic by containing all or some non-natural residues in place of naturally occurring amino acid residues. Nonnatural residues are well described in the scientific and patent literature; a few exemplary nonnatural compositions useful as mimetics of natural amino acid residues and guidelines are described below.

Mimetics of aromatic amino acids can be generated hy replacing by, e.g., D- or L- naphylalanine; D- or L- phenylglycine; D- or L-2 thieneylalanine; D- or L-1, -2, 3-, or 4-pyreneylalanine; D- or L-3 thieneylalanine; D- or L-(2-pyridinyl)-alanine; D- or L-(3-pyridinyl)-alanine; D- or L-(2-pyrazinyl)-alanine; D- or L-(4-isopropyl)-phenylglycine; D-(trifluoromethyl)-phenylglycine; D-(trifluoromethyl)-phenylalanine; D-p-fluorophenylalanine; D- or L-p-biphenylphenylalanine; K- or L-p-methoxyhiphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylainines, where alkyl can be substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, pentyl, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a nonnatural amino acid include, e.g., thiazolyl, thiophenyl, pyrazolyl, henzimidazolyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings.

Mimetics of acidic amino acids can he generated hy suhstitution hy, e.g., non-carboxylate amino acids while maintaining a negative charge; (phosphono)alanine; sulfated threonine. Carboxyl side groups (e.g., aspartyl or glutamyl) can also be selectively modified by reaction with carbodiimides (R=-N-C-N-R=) such as, e.g., 1-cyclohexyl-3(2-morpholinyl-

(4-ethyl) carbodiimide or 1-ethyl-3(4-azonia- 4,4- dimetholpentyl) carbodiimide. Aspartyl or glutamyl can also be converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Mimetics of basic amino acids can be generated by substitution with, e.g., (in addition to lysine and arginine) the amino acids ornithine, citrulline, or (guanidino)-acetic acid, or (guanidino)alkyl-acetic acid, where alkyl is defined above. Nitrile derivative (e.g., containing the CN-moiety in place of COOH) can be substituted for asparagine or glutamine. Asparaginyl and glutaminyl residues can be deaminated to the corresponding aspartyl or glutamyl residues.

Arginine residue mimetics can be generated by reacting arginyl with, e.g., one or more conventional reagents, including, e.g., phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, or ninhydrin, preferably under alkaline conditions.

Tyrosine residue mimetics can be generated by reacting tyrosyl with, e.g., aromatic diazonium compounds or tetranitromethane. N-acetylimidizol and tetranitromethane can be used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Cysteine residue mimetics can be generated by reacting cysteinyl residues with, e.g., alpha-haloacetates such as 2-chloroacetic acid or chloroacetamide and corresponding amines, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteine residue mimetics can also be generated by reacting cysteinyl residues with, e.g., bromo-trifluoroacetone, alpha-bromo-beta-(5-imidozoyl) propionic acid; chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide; methyl 2-pyridyl disulfide; p-chloromercuribenzoate; 2-chloromercuri-4 nitrophenol; or, chloro-7-nitrobenzo-oxa-1,3-diazole.

Lysine mimetics can be generated (and amino terminal residues can be altered) by reacting lysinyl with, e.g., succinic or other carboxylic acid anhydrides. Lysine and other alpha-amino-containing residue mimetics can also be generated by reaction with imidoesters, such as methyl picolinimidate, pyridoxal phosphate, pyridoxal, chloroborohydride, trinitrobenzenesulfonic acid, O-methylisourea, 2,4, pentanedione, and transamidase-catalyzed reactions with glyoxylate.

Mimetics of methionine can be generated by reaction with, e.g., methionine sulfoxide. Mimetics of proline include, e.g., pipecolic acid, thiazolidine carboxylic acid, 3- or 4- hydroxy proline, dehydroproline, 3- or 4-methylproline, or 3,3,-dimethylproline. Histidine residue mimetics can be generated by reacting histidyl with, e.g., diethylprocarbonate or para-

bromophenacyl bromide.

Other mimetics include, e.g., those generated by hydroxylation of proline and lysine; phosphorylation of the hydroxyl groups of seryl or threonyl residues; methylation of the alpha-amino groups of lysine, arginine and histidine; acetylation of the N-terminal amine; methylation of main chain amide residues or substitution with N-methyl amino acids; or amidation of C-terminal carboxyl groups.

A component of a natural polypeptide (e.g., a PL polypeptide or PDZ polypeptide) can also be replaced by an amino acid (or peptidomimetic residue) of the opposite chirality. Thus, any amino acid naturally occurring in the L-configuration (which can also be referred to as the R or S, depending upon the structure of the chemical entity) can be replaced with the amino acid of the same chemical structural type or a peptidomimetic, but of the opposite chirality, generally referred to as the D- amino acid, but which can additionally be referred to as the R- or S- form.

The mimetics of the invention can also include compositions that contain a structural mimetic residue, particularly a residue that induces or mimics secondary structures, such as a beta turn, beta sheet, alpha helix structures, gamma turns, and the like. For example, substitution of natural amino acid residues with D-amino acids; N-alpha-methyl amino acids; C-alpha-methyl amino acids; or dehydroamino acids within a peptide can induce or stabilize beta turns, gamma turns, beta sheets or alpha helix conformations. Beta turn mimetic structures have been described, e.g., by Nagai (1985) Tet. Lett. 26:647-650; Feigl (1986) J. Amer. Chem. Soc. 108:181-182; Kahn (1988) J. Amer. Chem. Soc. 110:1638-1639; Kemp (1988) Tet. Lett. 29:5057-5060; Kahn (1988) J. Molec. Recognition 1:75-79. Beta sheet mimetic structures have been described, e.g., by Smith (1992) J. Amer. Chem. Soc. 114:10672-10674. For example, a type VI beta turn induced by a cis amide surrogate, 1,5-disubstituted tetrazol, is described by Beusen (1995) Biopolymers 36:181-200. Incorporation of achiral omega-amino acid residues to generate polymethylene units as a substitution for amide bonds is described by Banerjee (1996) Biopolymers 39:769-777. Secondary structures of polypeptides can be analyzed by, e.g., high-field 1H NMR or 2D NMR spectroscopy, see, e.g., Higgins (1997) J. Pept. Res. 50:421-435. See also, Hruby (1997) Biopolymers 43:219-266, Balaji, et al., U.S. Pat. No. 5,612,895.

As used herein, "peptide variants" and "conservative amino acid substitutions"

refer to peptides that differ from a reference peptide (e.g., a peptide having the sequence of the carboxy-terminus of a specified PL protein) by substitution of an amino acid residue having similar properties (based on size, polarity, hydrophobicity, and the like). Thus, insofar as the compounds that are encompassed within the scope of the invention are partially defined in terms of amino acid residues of designated classes, the amino acids can be generally categorized into three main classes: hydrophilic amino acids, hydrophobic amino acids and cysteine-like amino acids, depending primarily on the characteristics of the amino acid side chain. These main classes may be further divided into subclasses. Hydrophilic amino acids include amino acids having acidic, basic or polar side chains and hydrophobic amino acids include amino acids having aromatic or apolar side chains. Apolar amino acids may be further subdivided to include, among others, aliphatic amino acids. The definitions of the classes of amino acids as used herein are as follows:

"Hydrophobic Amino Acid" refers to an amino acid having a side chain that is uncharged at physiological pH and that is repelled by aqueous solution. Examples of genetically encoded hydrophobic amino acids include Ile, Leu and Val. Examples of non-genetically encoded hydrophobic amino acids include t-BuA.

"Aromatic Amino Acid" refers to a hydrophobic amino acid having a side chain containing at least one ring having a conjugated π-electron system (aromatic group). The aromatic group may be further substituted with groups such as alkyl, alkenyl, alkynyl, hydroxyl, sulfanyl, nitro and amino groups, as well as others. Examples of genetically encoded aromatic amino acids include Phe, Tyr and Trp. Commonly encountered non-genetically encoded aromatic amino acids include phenylglycine, 2-naphthylalanine, β-2-thienylalanine, 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, 4-chloro-phenylalanine, 2-fluorophenylalanine, 3-fluorophenylalanine and 4-fluorophenylalanine.

"Apolar Amino Acid" refers to a hydrophobic amino acid having a side chain that is generally uncharged at physiological pH and that is not polar. Examples of genetically encoded apolar amino acids include Gly, Pro and Met. Examples of non-encoded apolar amino acids include Cha.

"Aliphatic Amino Acid" refers to an apolar amino acid having a saturated or unsaturated straight chain, branched or cyclic hydrocarbon side chain. Examples of genetically encoded aliphatic amino acids include Ala, Leu, Val and Ile. Examples of non-encoded aliphatic amino acids include Nle.

"Hydrophilic Amino Acid" refers to an amino acid having a side chain that is attracted by aqueous solution. Examples of genetically encoded hydrophilic amino acids include Ser and Lys. Examples of non-encoded hydrophilic amino acids include Cit and hCys.

"Acidic Amino Acid" refers to a hydrophilic amino acid having a side chain pK value of less than 7. Acidic amino acids typically have negatively charged side chains at physiological pH due to loss of a hydrogen ion. Examples of genetically encoded acidic amino acids include Asp and Glu.

"Basic Amino Acid" refers to a hydrophilic amino acid having a side chain pK value of greater than 7. Basic amino acids typically have positively charged side chains at physiological pH due to association with hydronium ion. Examples of genetically encoded basic amino acids include Arg, Lys and His. Examples of non-genetically encoded basic amino acids include the non-cyclic amino acids ornithine, 2,3-diaminopropionic acid, 2,4-diaminobutyric acid and homoarginine.

"Polar Amino Acid" refers to a hydrophilic amino acid having a side chain that is uncharged at physiological pH, but which has a bond in which the pair of electrons shared in common by two atoms is held more closely by one of the atoms. Examples of genetically encoded polar amino acids include Asx and Glx. Examples of non-genetically encoded polar amino acids include citrulline, N-acetyl lysine and methionine sulfoxide.

"Cysteine-Like Amino Acid" refers to an amino acid having a side chain capable of forming a covalent linkage with a side chain of another amino acid residue, such as a disulfide linkage. Typically, cysteine-like amino acids generally have a side chain containing at least one thiol (SH) group. Examples of genetically encoded cysteine-like amino acids include Cys. Examples of non-genetically encoded cysteine-like amino acids include homocysteine and penicillamine.

As will be appreciated by those having skill in the art, the above classification are not absolute -- several amino acids exhibit more than one characteristic property, and can therefore be included in more than one category. For example, tyrosine has both an aromatic ring and a polar hydroxyl group. Thus, tyrosine has dual properties and can be included in both the aromatic and polar categories. Similarly, in addition to being able to form disulfide linkages, cysteine also has apolar character. Thus, while not strictly classified as a hydrophobic or apolar amino acid, in many instances cysteine can be used to confer hydrophobicity to a peptide.

Certain commonly encountered amino acids which are not genetically encoded of which the peptides and peptide analogues of the invention are composed include, but are not limited to, β-alanine (b-Ala) and other omega-amino acids such as 3-aminopropionic acid (Dap), 2,3-diaminopropionic acid (Dpr), 4-aminobutyric acid and so forth; α-aminoisobutyric acid (Aib); ε-aminohexanoic acid (Aha); δ-aminovaleric acid (Ava); N-methylglycine or sarcosine (MeGly); ornithine (Orn); citrulline (Cit); t-butylalanine (t-BuA); t-butylglycine (t-BuG); N-methylisoleucine (MeIle); phenylglycine (Phg); cyclohexylalanine (Cha); norleucine (Nle); 2-naphthylalanine (2-Nal); 4-chlorophenylalanine (Phe(4-Cl)); 2-fluorophenylalanine (Phe(2-F)); 3-fluorophenylalanine (Phe(3-F)); 4-fluorophenylalanine (Phe(4-F)); penicillamine (Pen); 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Tic); β-2-thienylalanine (Thi); methionine sulfoxide (MSO); homoarginine (hArg); N-acetyl lysine (AcLys); 2,3-diaminobutyric acid (Dab); 2,3-diaminobutyric acid (Dbu); p-aminophenylalanine (Phe(pNH_2)); N-methyl valine (MeVal); homocysteine (hCys) and homoserine (hSer). These amino acids also fall conveniently into the categories defined above.

The classifications of the above-described genetically encoded and non-encoded amino acids are summarized in **TABLE 1**, below. It is to be understood that **TABLE 1** is for illustrative purposes only and does not purport to be an exhaustive list of amino acid residues which can comprise the peptides and peptide analogues described herein. Other amino acid residues which are useful for making the peptides and peptide analogues described herein can be found, e.g., in Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology, CRC Press, Inc., and the references cited therein. Amino acids not specifically mentioned herein can be conveniently classified into the above-described categories on the basis of known behavior and/or their characteristic chemical and/or physical properties as compared with amino acids specifically identified.

TABLE 1

Classification	Genetically Encoded	Genetically Non-Encoded
Hydrophobic		
Aromatic	F, Y, W	Phg, Nal, Thi, Tic, Phe(4-Cl), Phe(2-F), Phe(3-F), Phe(4-F), Pyridyl Ala, Benzothienyl Ala
Apolar	M, G, P	
Aliphatic	A, V, L, I	t-BuA, t-BuG, MeIle, Nle, MeVal, Cha, bAla, MeGly, Aib

Hydrophilic		
Acidic	D, E	
Basic	H, K, R	Dpr, Orn, hArg, Phe(p-NH_2), DBU, A_2BU
Polar	Q, N, S, T, Y	Cit, AcLys, MSO, hSer
Cysteine-Like	C	Pen, hCys, p-methyl Cys

As used herein, a "detectable label" has the ordinary meaning in the art and refers to an atom (e.g., radionuclide), molecule (e.g., fluorescein), or complex, that is or can be used to detect (e.g., due to a physical or chemical property), indicate the presence of a molecule or to enable binding of another molecule to which it is covalently bound or otherwise associated. The term "label" also refers to covalently bound or otherwise associated molecules (e.g., a biomolecule such as an enzyme) that act on a substrate to produce a detectable atom, molecule or complex. Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Labels useful in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (e.g., DynabeadsTM), fluorescent dyes (e.g., fluorescein, Texas red, rhodamine, green fluorescent protein, enhanced green fluorescent protein, and the like), radiolabels (e.g., ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (e.g., hydrolases, particularly phosphatases such as alkaline phosphatase, esterases and glycosidases, or oxidoreductases, particularly peroxidases such as horse radish peroxidase, and others commonly used in ELISAs), substrates, cofactors, inhibitors, chemiluminescent groups, chromogenic agents, and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Means of detecting such labels are well known to those of skill in the art. Thus, for example, radiolabels and chemiluminescent labels can be detected using photographic film or scintillation counters, fluorescent markers can be detected using a photodetector to detect emitted light (e.g., as in fluorescence-activated cell sorting). Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label. Thus, a label is any composition detectable by

spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. The label can be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (e.g., streptavidin) molecule which is either inherently detectable or covalently bound to a signal generating system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody. The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter, photographic film as in autoradiography, or storage phosphor imaging. Where the label is a fluorescent label, it can be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence can be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels can be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Also, simple colorimetric labels can be detected by observing the color associated with the label. It will be appreciated that when pairs of fluorophores are used in an assay, it is often preferred that they have distinct emission patterns (wavelengths) so that they can be easily distinguished.

As used herein, the term "substantially identical" in the context of comparing amino acid sequences, means that the sequences have at least about 70%, at least about 80%, or at least about 90% amino acid residue identity when compared and aligned for maximum correspondence. An algorithm that is suitable for determining percent sequence identity and sequence similarity is the FASTA algorithm, which is described in Pearson, W.R. & Lipman, D.J., 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85: 2444. See also *W. R. Pearson, 1996, Methods Enzymol.* 266: 227-258. Preferred parameters used in a FASTA alignment of DNA sequences to calculate percent identity are optimized, BL50 Matrix 15: -5, k-tuple = 2; joining penalty =

40, optimization = 28; gap penalty -12, gap length penalty =-2; and width = 16.

As used herein, "hematopoietic cells" include leukocytes including lymphocytes (T cells, B cells and NK cells), monocytes, and granulocytes (i.e., neutrophils, basophils and eosinophils), macrophages, dendritic cells, megakaryocytes, reticulocytes, erythrocytes, and CD34$^+$ stem cells.

As used herein, the terms "test compound" or "test agent" are used interchangeably and refer to a candidate agent that can have enhancer/agonist, or inhibitor/antagonist activity, e.g., inhibiting or enhancing an interaction such as PDZ-PL binding. The candidate agents or test compounds can be any of a large variety of compounds, both naturally occurring and synthetic, organic and inorganic, and including polymers (e.g., oligopeptides, polypeptides, oligonucleotides, and polynucleotides), small molecules, antibodies (as broadly defined herein), sugars, fatty acids, nucleotides and nucleotide analogs, analogs of naturally occurring structures (e.g., peptide mimetics, nucleic acid analogs, and the like), and numerous other compounds. In certain embodiment, test agents are prepared from diversity libraries, such as random or combinatorial peptide or non-peptide libraries. Many libraries are known in the art that can be used, *e.g.*, chemically synthesized libraries, recombinant (*e.g.*, phage display libraries), and *in vitro* translation-based libraries. Examples of chemically synthesized libraries are described in Fodor et al., 1991, *Science* 251:767-773; Houghten et al., 1991, *Nature* 354:84-86; Lam et al., 1991, *Nature* 354:82-84; Medynski, 1994, *Bio/Technology* 12:709-710; Gallop et al., 1994, J. *Medicinal Chemistry* 37(9):1233-1251; Ohlmeyer et al., 1993, *Proc. Natl. Acad. Sci. USA* 90:10922-10926; Erb et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:11422-11426; Houghten et al., 1992, *Biotechniques* 13:412; Jayawickreme et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:1614-1618; Salmon et al., 1993, *Proc. Natl. Acad. Sci. USA* 90:11708-11712; PCT Publication No. WO 93/20242; and Brenner and Lerner, 1992, *Proc. Natl. Acad. Sci. USA* 89:5381-5383. Examples of phage display libraries are described in Scott and Smith, 1990, *Science* 249:386-390; Devlin et al., 1990, *Science*, 249:404-406; Christian, R.B., et al., 1992, *J. Mol. Biol.* 227:711-718); Lenstra, 1992, *J. Immunol. Meth.* 152:149-157; Kay et al., 1993, *Gene* 128:59-65; and PCT Publication No. WO 94/18318 dated August 18, 1994. *In vitro* translation-based libraries include but are not limited to those described in PCT Publication No. WO 91/05058 dated April 18, 1991; and

Mattheakis et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:9022-9026. By way of examples of nonpeptide libraries, a benzodiazepine library (*see e.g.*, Bunin et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:4708-4712) can be adapted for use. Peptoid libraries (Simon et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:9367-9371) can also be used. Another example of a library that can be used, in which the amide functionalities in peptides have been permethylated to generate a chemically transformed combinatorial library, is described by Ostresh et al. (1994, *Proc. Natl. Acad. Sci. USA* 91:11138-11142).

The term "specific binding" refers to binding between two molecules, for example, a ligand and a receptor, characterized by the ability of a molecule (ligand) to associate with another specific molecule (receptor) even in the presence of many other diverse molecules, i.e., to show preferential binding of one molecule for another in a heterogeneous mixture of molecules. Specific binding of a ligand to a receptor is also evidenced by reduced binding of a detectably labeled ligand to the receptor in the presence of excess unlabeled ligand (i.e., a binding competition assay).

As used herein, a "plurality" of PDZ proteins (or corresponding PDZ domains or PDZ fusion polypeptides) has its usual meaning. In some embodiments, the plurality is at least 5, and often at least 25, at least 40, or at least 60 different PDZ proteins. In some embodiments, the plurality is selected from the list of PDZ polypeptides listed in **Table 9**. In some embodiments, the plurality of different PDZ proteins are from (i.e., expressed in) a particular specified tissue or a particular class or type of cell. In some embodiments, the plurality of different PDZ proteins represents a substantial fraction (e.g., typically at least 50%, more often at least 80%) of all of the PDZ proteins known to be, or suspected of being, expressed in the tissue or cell(s), e.g., all of the PDZ proteins known to be present in lymphocytes or hematopoetic cells. In some embodiments, the plurality is at least 50%, usually at least 80%, at least 90% or all of the PDZ proteins disclosed herein as being expressed in a particular cell.

When referring to PL peptides (or the corresponding proteins, e.g., corresponding to those listed in **TABLE 8**, or elsewhere herein) a "plurality" can refer to at least 5, at least 10, and often at least 25 PLs such as those specifcally listed herein, or to the classes and percentages set forth *supra* for PDZ domains.

II. Overview

The present inventors have identified a number of interactions between PDZ proteins and PL proteins that can play a significant role in the biological function of certain cells and in a variety of physiological systems. As used herein, the term "biological function" in the context of a cell, refers to a detectable biological activity normally carried out by the cell, e.g., a phenotypic change such as proliferation, cell activation (e.g., T cell activation, B cell activation, T-B cell conjugate formation), cytokine release, degranulation, tyrosine phosphorylation, ion (e.g., calcium) flux, metabolic activity, apoptosis, changes in gene expression, maintenance of cell structure, cell migration, adherence to a substrate, signal transduction, cell-cell interactions, and others described herein or known in the art.

Because the interactions involve proteins that are involved in diverse physiological systems, the methods provided herein can be utilized broadly or selectively to modulate a number of different biological functions. Methods are also disclosed herein for determining whether a test compound acts as a modulator of binding between a particular PDZ protein and PL protein binding pair. Both agonists and antagonists of the binding pairs can be identified by such screening methods. Modulators so identified can subsequently be formulated as a pharmaceutical composition and used in the treatment of various diseases that are correlated with binding between a particular PDZ protein and PL protein or set of such proteins.

III. PDZ Protein and PL Protein Interactions

TABLE 7 and **TABLE 12** (located at the end of the specification) list PDZ proteins and PL proteins which the current inventors have identified as binding to one another using assay methods described *infra*. Each page of **TABLE 7** and **12** includes seven columns. The columns in each table are numbered from left to right such that the left-most column in each table is column 1 and the right-most column in each table is column 7. Thus, the first column in each table is labeled "AVC ID"; this column simply lists an internal reference number used to refer to the carboxyl-terminal amino acids of the PL proteins listed in the second column. Thus, the second column labeled "PL" lists the various PL peptides that were identified as binding a PDZ protein. All PL peptides were biotinylated at the amino-terminus and the sequences of these PL peptides are presented in **TABLE 8** (see end of specification).

The PDZ protein (or proteins) that interact(s) with a PL peptide are listed in the fourth column of each table that is labeled "PDZ". This column provides the gene name for the PDZ portion of the GST-PDZ fusion that interacts with the PDZ-ligand to the left. For PDZ domain-containing proteins with multiple domains, the domain number is listed to the right of the PDZ (i.e., in column 5 labeled "PDZ Domain"), and indicates the PDZ domain number when numbered from the amino-terminus to the carboxy-terminus.

The third column labeled "*Peptide* Optimal Concentration" in the tables is a number reflective of the binding interaction between the PL protein and PDZ protein. If a '0' is listed, this is an indication that an interaction was observed using a PL peptide concentration of 10 uM in the assay; any other value listed is indicative of the Kd (dissociation constant) in uM determined by titration of the PL peptide on the concentration of PDZ protein listed in **TABLE** 7 and 12 (see infra for methods for determining Kd). The column labeled "*Protein* Optimal Concentration" refers to the protein concentration used to assay PL interaction (in ug/ml); a '0' is indicative of 5 ug/ml protein concentration; any other value represents the concentration (in ug/ml) used to determine the dissociation constant for a given interaction.

Finally, the seventh column labeled "Classification" is another measure of the level of binding. In particular, it provides an absorbance value at 450 nm which indicates the amount of PL peptide bound to the PDZ protein. The following numerical values have the following meanings: '1' – A_{450}nm 0-1; '2' - A_{450}nm 1-2; '3'- A_{450}nm 2-3; '4' - A_{450}nm 3-4; '5' - A_{450}nm of 4 more than 2X repeated; '0' – A_{450}nm 0, i.e., not found to interact. Thus, higher numbers indicate stronger interactions.

Further information regarding these PL proteins and PDZ proteins is provided in **TABLES** 8 and 9. In particular, TABLE 8 provides a listing of the amino acid sequences of peptides used in the assays. When numbered from left to right, the first column labeled "AVC ID" provides the internal designation number used to refer to a particular PL protein and correlates with the designation used in TABLE 7 or TABLE 12. The column labeled "AVC Name" provides the name of the gene containing a predicted PDZ ligand. The third column labeled "Sequence" is the amino acid sequence of the PL protein used in the assay. The final two columns labeled "Accession No. and GI list the Genbank accession number or GI number corresponding to the sequence and gene name. As indicated supra, all peptides are biotinylated at the amino terminus and the amino acid sequences correspond to the C-terminal sequence of the gene name listed to the left.

TABLE 9 (located at the end of the specification) lists the sequences of the PDZ domains cloned into a vector (PGEX-3X vector) for production of GST-PDZ fusion proteins (Pharmacia) (see section VI (A)) below). More specifically, the first column (left to right) entitled "Gene Name" lists the name of the gene containing the PDZ domain. The second column labeled "GI" is a unique Genbank identifier for the gene used to design primers for PCR amplification of the listed sequence. The next column labeled "Domain Number" indicates the Pfam-predicted PDZ domain number, as numbered from the amino-terminus of the gene to the carboxy-terminus. The last column entitled "Sequence" provides the actual amino acid sequence inserted into the GST-PDZ expression vector as determined by DNA sequencing of the constructs.

As discussed in detail herein, the PDZ proteins listed in **TABLE 7** and **12** are naturally occurring proteins containing a PDZ domain. Only significant interactions are presented in **TABLE 7** and **12.** Thus, the present invention is particularly directed to the detection and modulation of interactions between a PDZ protein and PL protein pair listed in **TABLE 7** or in **12.** As used herein the phrase "protein pair" or 'protein binding pair" when used in reference to a PDZ protein and PL protein refers to a PL protein and PDZ protein listed in **TABLE 7** or **12** which bind to one another. It should be understood that **TABLE 7** and **12** are set up to show that certain PL proteins bind to a plurality of PDZ proteins. For example, in **TABLE 7,** PL protein CD46 (page 2 of TABLE 2) binds to the PDZ proteins KIAA0973, Mint 1, KIAA807, BAI-1, KIA0807(S), and PL protein CX43 binds to PDZ proteins ZO-2 andZO-1.

IV. Classification of Interactions

A. General

The interactions summarized in **TABLE 7** and **12** can occur in a wide variety of cell types. Examples of such cells include hematopoietic, stem, neuronal, muscle, epidermal, epithelial, endothelial, and cells from essentially any tissue such as liver, lung, placenta, uterus, kidney, ovaries, testes, stomach, colon and intestine. Because the interactions disclosed herein can occur in such a wide variety of cell types, these interactions can also play an important role in a variety of biological functions. Consequently, modulation of the interactions between PDZ proteins and PL proteins that are described herein can be utilized to regulate biological function in a wide range of cells.

In certain methods disclosed herein, the PL protein is expressed or up-regulated upon cell activation (e.g., in activated B lymphocytes, T lymphocytes) or upon entry into mitosis (e.g., up-regulation in rapidly proliferating cell populations).

B. Exemplary PDZ Classification

The PDZ proteins identified herein as interacting with particular PL proteins can be grouped into a number of different categories. Thus, as described in greater detail below, the methods and reagents that are provided herein can be utilized to modulate PDZ interactions, and thus biological functions, that are regulated or otherwise involve the following classes of proteins. It should be recognized, however, that modulation of the interactions that are identified herein can be utilized to affect biological functions involving other protein classes.

1. Protein Kinases

A number of protein kinases contain PDZ domains. Protein kinases are widely involved in cellular metabolism and regulation of protein activity through phosphorylation of amino acids on proteins. An example of this the regulation of signal transduction pathways such as T cell activation throuh the T cell Receptor, where ZAP-70 kinase function is required for transmission of the activation signal to downstream effector molecules. These molecules include, but are not limited to KIAA0303, KIAA0561, KIAA0807, KIAA0973, and CASK.

2. Guanalyte Kinases

A number of guanalyte kinases contain PDZ domains. These molecules include, but are not limited to Atrophin 1, CARD11, CARD14, DLG1, DLG2, DLG5, FLJ12615, MPP1, MPP2, NeDLG, p55T, PSD95, ZO-1, ZO-2, and ZO-3.

3. Guanine Exchange Factors

A number of guanine exchange factors contain PDZ domains. Guanine exchange factors regulate signal transduction pathways and other biological processes through facilitating the exchange of differentially phosphorylated guanine residues. These molecules include, but are not limited to GTPase, Guanine Exchange, KIAA0313,

KIAA0380, KIAA0382, KIAA1389, KIAA1415, TIAM1, and TIAM2.

4. LIM PDZ's

A number of LIM PDZ's contain PDZ domains. These molecules include, but are not limited to Alpha Actinin 2, ELFIN1, ENIGMA, HEMBA 1003117, KIAA0613, KIAA0858, KIAA0631, LIM Mystique, LIM protein, LIM-RIL, LIMK1, LIMK2, and LU-1.

5. Proteins Containing Only PDZ Domains

A number of proteins contain PDZ domains without any other predicted functional domains. These include, but are not limited to 26s subunit p27, AIPC, Cytohesin Binding, EZRIN Binding Protein, FLJ00011, FLJ20075, FLJ21687, GRIP1, HEMBA1000505, KIAA0545, KIAA0967, KIAA1202, KIAA1284, KIAA1526, KIAA1620, KIAA1719, MAGGI1; Novel PDZ gene, Outer Membrane, PAR3, PAR6, PAR6 Gamma, PDZ-73, PDZK1, PICK1, PIST, prIL16, Shank 1, SIP1, SITAC-18, SYNTENIN, Syntrophin gamma 2, TIP1, TIP2, and TIP43.

6. Tyrosine Phosphatases

A number of Tyrosine phosphatases contain PDZ domains. Tyrosine phosphatases regulate biological processes such as signal transduction pathways through removal of phosphate groups required for function of the target protein. Examples of such enzymes include, but are not limited to PTN-3, PNT-4, and PTPL1.

7. Serine Proteases

A number of Serine Proteases contain PDZ domains. Proteases affect biological molecules by cleaving them to either activate or repress their functional ability. These enzymes have a variety of functions, including roles in digestion, blood coagulation and lysis of blood clots. These include, but are not limited to Novel Serine Protease, and Serine Protease.

8. Viral Oncogene Interacting Proteins That Contain PDZ Domains

A number of TAX interacting proteins contain PDZ domains. Many of these

also bind to multiple viral oncoproteins such as Adenovirus E4, Papillomavirus E6, and HBV protein X. These include, but are not limited to AIPC, Connector Enhancer, DLG1, DLG2, ERBIN, FLJ00011, FLJ11215, HEMBA 1003117, INADL, KIAA0147, KIAA0807, KIAA1526, KIAA1634, LIMK1, LIM Mystique, LIM-RIL, MUPP1, NeDLG, Outer Membrane, PSD95, PTN-4, PTPL1, Syntrophin gamma 1, Syntrophin gamma 2, TAX2-like protein, TIP2, TIP1, TIP33 and TIP43.

9. Proteins Containing RA and/or RHA and/or DIL and/or IGFBP and/or WW and/or L27 and/or SAM and/or PH and/or DIX and/or DIP and/or Dishevelled and/or LRR and/or Hormone 3 and/or C2 and/or RPH3A and/or zf-TRAF and/or zf-C3HC4 and/or PID and/or NO_Synthase and/or Flavodoxin and/or FAD binding and/or NAD binding and/or Kazal, and/or Trypsin an/or RBD and/or RGS and/or GoLoco and/orHR1 and/or BR01 That Contain PDZ Domains

A number o proteins containing RA and/or RHA and/or DIL and/or IGFBP and/or WW and/or L27 and/or SAM and/or PH and/or DIX and/or DIP and/or Dishevelled and/or LRR and/or Hormone 3 and/or C2 and/or RPH3A and/or zf-TRAF and/or zf-C3HC4 and/or PID and/or NO_Synthase and/or Flavodoxin and/or FAD binding and/or NAD binding and/or Kazal, and/or Trypsin an/or RBD and/or RGS and/or GoLoco and/orHR1 and/or BR01 contain PDZ domains. These include, but are not limited to AF6, APXL-1, BAI-1 Associated, DVL1, DVL2, DVL3, KIAA0417, KIAA0316, KIAA0340, KIAA0559, KIAA0751, KIAA0902, KIAA1095, KIA1222, KIAA1634, MINT1, NOS1, RGS12, Rhophilin-like, Shank3, Syntrophin 1 alpha, Syntrophin beta 2, and X-11 beta.

C. Exemplary PL Classification

The PL proteins involved in the interactions listed in **TABLE** 7 and 12 are from a number of different classes. Consequently, the methods and reagents that are disclosed herein can be utilized to to modulate interactions involving the following classes of PL proteins to modulate a bioloigcal function in cells. The following classes, however, should not be considered exhaustive of the the types of classes of proteins whose activity can be modulated using the methods and reagents that are provided herein.

1. PL Sequences of T Cell Surface Receptors

A number of surface receptors expressed by T cells contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, CD6, CD95, CDw128B (IL8 R), DNAM-1, Fas ligand (FasL), LPAP (Barclay et al., 1997, The Leucocyte Antigen Facts Book, second edition, Academic Press), CLASP-1, CLASP-2, CLASP-5, BLR-1 (CXCR5), DOCK2, PAG, and Mannose Receptor.

The C-terminal core sequence of CD6 is ISAA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AA (SEQ ID NO:X), SAA (SEQ ID NO:X), DISAA (SEQ ID NO:X), DDISAA (SEQ ID NO:X), YDDISAA (SEQ ID NO:X), and DYDDISAA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of CD95 is QSLV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LV (SEQ ID NO:X), SLV (SEQ ID NO:X), IQSLV (SEQ ID NO:X), EIQSLV (SEQ ID NO:X), NEIQSLV (SEQ ID NO:X), and RNEIQSLV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of CDw128B is STTL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, TL (SEQ ID NO:X), TTL (SEQ ID NO:X), TSTTL (SEQ ID NO:X), HTSTTL (SEQ ID NO:X), GHTSTTL (SEQ ID NO:X), and SGHTSTTL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of DNAM-1 is KTRV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RV (SEQ ID NO:X), TRV (SEQ ID NO:X), PKTRV (SEQ ID NO:X), RPKTRV (SEQ ID NO:X), RRPKTRV (SEQ ID NO:X), and SRRPKTRV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of FasL is LYKL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KL (SEQ ID NO:X), YKL (SEQ ID NO:X), GLYKL (SEQ ID NO:X), FGLYKL (SEQ ID NO:X), FFGLYKL (SEQ ID NO:X), and TFFGLYKL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of LPAP is VTAL (SEQ ID NO:X). When

naturally-occurring residues are added or removed from the core sequence, AL (SEQ ID NO:X), TAL (SEQ ID NO:X), HVTAL (SEQ ID NO:X), LHVTAL (SEQ ID NO:X), GLHVTAL (SEQ ID NO:X), and QGLHVTAL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of CLASP-1 is SAQV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, QV (SEQ ID NO:X), AQV (SEQ ID NO:X), SSAQV (SEQ ID NO:X), SSSAQV (SEQ ID NO:X), ISSSAQV (SEQ ID NO:X), and SISSSAQV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of CLASP-2 is SSVV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VV (SEQ ID NO:X), SVV (SEQ ID NO:X), SSSVV (SEQ ID NO:X), SSSSVV (SEQ ID NO:X), TSSSSVV (SEQ ID NO:X), and MTSSSSVV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of CLASP-5 is SQGS (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, GS (SEQ ID NO:X), QGS (SEQ ID NO:X), LSQGS (SEQ ID NO:X), QLSQGS (SEQ ID NO:X), TQLSQGS (SEQ ID NO:X), and ETQLSQGS (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of BLR-1 is LTTF (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, TF (SEQ ID NO:X), TTF (SEQ ID NO:X), SLTTF (SEQ ID NO:X), TSLTTF (SEQ ID NO:X), ATSLTTF (SEQ ID NO:X), and NATSLTTF (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of DOCK2 is STDL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DL (SEQ ID NO:X), TDL (SEQ ID NO:X), LSTDL (SEQ ID NO:X), SLSTDL (SEQ ID NO:X), DSLSTDL (SEQ ID NO:X), and PDSLSTDL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of PAG is ITRL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RL (SEQ ID NO:X), TRL (SEQ ID NO:X), DITRL (SEQ ID NO:X), RDITRL (SEQ ID NO:X),

GRDITRL (SEQ ID NO:X), and QGRDITRL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

The C-terminal core sequence of Mannose Receptor is HSVI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VI (SEQ ID NO:X), SVI (SEQ ID NO:X), EHSVI (SEQ ID NO:X), NEHSVI (SEQ ID NO:X), QNEHSVI (SEQ ID NO:X), and EQNEHSVI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in T cells.

2. PL Sequences of B Cell Surface Receptors

A number of surface receptors expressed by B cells contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, CD95, CDW125 (modified) (IL5R), DNAM-1, LPAP (Barclay et al., 1997, The Leucocyte Antigen Facts Book, second edition, Academic Press), CLASP-1, CLASP-2, CLASP-5, and BLR-1. The specific motif sequences of CD95, DNAM-1, LPAP, CLASP-1, CLASP-2, CLASP-5, and BLR-1 have been described in the preceding paragraphs.

The C-terminal core sequence of CDW125 is DSVF (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VF (SEQ ID NO:X), SVF (SEQ ID NO:X), EDSVF (SEQ ID NO:X), LEDSVF (SEQ ID NO:X), TLEDSVF (SEQ ID NO:X), and ETLEDSVF (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in B cells.

3. PL Sequences of Natural Killer Cell Surface Receptors

A number of surface receptors expressed by NK cells contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, DNAM1. The specific motif sequence of DNAM-1 has been described in the preceding paragraphs.

4. PL Sequences of Monocyte Surface Receptors

A number of surface receptors expressed by cells of the monocytic lineage (monocytes and macrophages) contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, CD46, CD95, CDw128, DNAM-1, Mannose receptor, and FcεRIβ. The specific motif sequences of CD95, CDw128B, DNAM-1, and Mannose receptor have been described in the preceding paragraphs.

The C-terminal core sequence of CD46 is FTSL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SL (SEQ ID NO:X), TSL (SEQ ID NO:X), KFTSL (SEQ ID NO:X), VKFTSL (SEQ ID NO:X), EVKFTSL (SEQ ID NO:X), and REVKFTSL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in monocytes.

The C-terminal core sequence of FcεRIβ is PIDL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DL (SEQ ID NO:X), IDL (SEQ ID NO:X), PPIDL (SEQ ID NO:X), SPPIDL (SEQ ID NO:X), MSPPIDL (SEQ ID NO:X), and EMSPPIDL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in monocytes.

5. PL Sequences of Granulocyte Surface Receptors

A number of surface receptors expressed by granulocytes contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, CD95, CDW125, and FcεRIβ. The specific motif sequences of CD95, CDW125, and FcεRIβ have been described in the preceding paragraphs.

6. PL Sequences of Endothelial Cell Surface Receptors

A number of surface receptors expressed by endothelial cells contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, CD34, and CD46. The specific motif sequence of CD46 has been described in the preceding paragraphs.

The C-terminal core sequence of CD34 is DTEL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EL (SEQ ID NO:X), TEL (SEQ ID NO:X), ADTEL (SEQ ID NO:X), VADTEL (SEQ ID NO:X), VVADTEL (SEQ ID NO:X), and HVVADTEL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in endothelial cells.

7. PL Sequences of G-Protein Linked Receptors

A number of G-protein linked receptors contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, alpha-2A Adrenergic receptor, alpha-2B Adrenergic receptor, alpha-2C Adrenergic receptor, GLUR2, GluR5-2 (rat),

GLUR7, GluR delta-2, muscarinic Ach receptor M4, NMDA Glutamate Receptor 2C (cysteine-free), NMDA R2C, Serotonin receptor 3a, serotonin receptor 5-HT-2B, serotonin receptor 5-HT-2C, SSTR2 (somatostatin receptor 2), somatostatin receptor 4, IL-8RA, parathyroid hormone receptor 2, and C5 Anaphylatoxin receptor.

The C-terminal core sequence of alpha-2A Adrenergic receptor is KRIV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, IV (SEQ ID NO:X), RIV (SEQ ID NO:X), RKRIV (SEQ ID NO:X), DRKRIV (SEQ ID NO:X), GDRKRIV (SEQ ID NO:X), and RGDRKRIV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of alpha-2B Adrenergic receptor is QTAW (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AW (SEQ ID NO:X), TAW (SEQ ID NO:X), TQTAW (SEQ ID NO:X), WTQTAW (SEQ ID NO:X), PWTQTAW (SEQ ID NO:X), and RPWTQTAW (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of alpha-2C Adrenergic receptor is GFRQ (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RQ (SEQ ID NO:X), FRQ (SEQ ID NO:X), RGFRQ (SEQ ID NO:X), RRGFRQ (SEQ ID NO:X), ARRGFRQ (SEQ ID NO:X), and RARRGFRQ (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GLUR2 is SVKI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KI (SEQ ID NO:X), VKI (SEQ ID NO:X), ESVKI (SEQ ID NO:X), IESVKI (SEQ ID NO:X), GIESVKI (SEQ ID NO:X), and SGIESVKI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GLUR5-2 is ETVA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VA (SEQ ID NO:X), TVA (SEQ ID NO:X), KETVA (SEQ ID NO:X), RKETVA (SEQ ID NO:X), QRKETVA (SEQ ID NO:X), and TQRKETVA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GLUR7 is NLVI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VI (SEQ ID NO:X), LVI (SEQ ID NO:X), NNLVI (SEQ ID NO:X), YNNLVI (SEQ ID NO:X),

SYNNLVI (SEQ ID NO:X), and VSYNNLVI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GluR delta-2 is GTSI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SI (SEQ ID NO:X), TSI (SEQ ID NO:X), RGTSI (SEQ ID NO:X), DRGTSI (SEQ ID NO:X), PDRGTSI (SEQ ID NO:X), and DPDRGTSI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of muscarinic Ach receptor M4 is EQAL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AL (SEQ ID NO:X), QAL (SEQ ID NO:X), PEQAL (SEQ ID NO:X), APEQAL (SEQ ID NO:X), RAPEQAL (SEQ ID NO:X), and KRAPEQAL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of NMDA Glutamate Receptor 2C is ESEV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EV (SEQ ID NO:X), SEV (SEQ ID NO:X), LESEV (SEQ ID NO:X), SLESEV (SEQ ID NO:X), SSLESEV (SEQ ID NO:X), and ISSLESEV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of NMDA R2C is STVV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VV (SEQ ID NO:X), TVV (SEQ ID NO:X), VSTVV (SEQ ID NO:X), SVSTVV (SEQ ID NO:X), PSVSTVV (SEQ ID NO:X), and DPSVSTVV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Serotonin receptor 3a is WQYA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YA (SEQ ID NO:X), QYA (SEQ ID NO:X), IWQYA (SEQ ID NO:X), SIWQYA (SEQ ID NO:X), WSIWQYA (SEQ ID NO:X), and LWSIWQYA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of serotonin receptor 5-HT-2B is VSYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), SYV (SEQ ID NO:X), QVSYV (SEQ ID NO:X), EQVSYV (SEQ ID NO:X), EEQVSYV (SEQ ID NO:X), and TEEQVSYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of serotonin receptor 5-HT-2C is ISSV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SV (SEQ ID NO:X), SSV (SEQ ID NO:X), RISSV (SEQ ID NO:X), ERISSV (SEQ ID NO:X), SERISSV (SEQ ID NO:X), and VSERISSV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of SSTR 2 is QTSI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SI (SEQ ID NO:X), TSI (SEQ ID NO:X), LQTSI (SEQ ID NO:X), DLQTSI (SEQ ID NO:X), GDLQTSI (SEQ ID NO:X), and NGDLQTSI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of somatostatin receptor 4 is TTTF (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, TF (SEQ ID NO:X), TTF (SEQ ID NO:X), RTTTF (SEQ ID NO:X), TRTTTF (SEQ ID NO:X), LTRTTTF (SEQ ID NO:X), and PLTRTTTF (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of IL-8RA is SSNL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, NL (SEQ ID NO:X), SNL (SEQ ID NO:X), VSSNL (SEQ ID NO:X), NVSSNL (SEQ ID NO:X), VNVSSNL (SEQ ID NO:X), and SVNVSSNL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of parathyroid hormone receptor 2 is EDVL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VL (SEQ ID NO:X), DVL (SEQ ID NO:X), TEDVL (SEQ ID NO:X), ETEDVL (SEQ ID NO:X), GETEDVL (SEQ ID NO:X), and QGETEDVL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of C5 Anaphylatoxin receptor is TQAV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AV (SEQ ID NO:X), QAV (SEQ ID NO:X), KTQAV (SEQ ID NO:X), QKTQAV (SEQ ID NO:X), AQKTQAV (SEQ ID NO:X), and MAQKTQAV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

8. PL Sequences of Viral Oncogenes

A number of viral oncogenes and viral oncogene homologues contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, AdenoE4 typ9, AKT1, HPV E6 #16 (Modified), HPV E6 #18, HPV E6 33 (modified), HPV E6 #35 (cysteine-free), HPV E6 52 (modified), HPV E6 #57 (cysteine-free), HPV E6 58 (modified), HPV E6 #66 (cysteine-free), HPV E6 77 (modified), and TAX.

The C-terminal core sequence of AdenoE4 typ9 is ATLV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LV (SEQ ID NO:X), TLV (SEQ ID NO:X), IATLV (SEQ ID NO:X), KIATLV (SEQ ID NO:X), VKIATLV (SEQ ID NO:X), and SVKIATLV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of AKT1 is SSTA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, TA (SEQ ID NO:X), STA (SEQ ID NO:X), ASSTA (SEQ ID NO:X), SASSTA (SEQ ID NO:X), YSASSTA (SEQ ID NO:X), and SYSASSTA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 #16 is ETQL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, QL (SEQ ID NO:X), TQL (SEQ ID NO:X), RETQL (SEQ ID NO:X), RRETQL (SEQ ID NO:X), TRRETQL (SEQ ID NO:X), and RTRRETQL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 #18 is ETQV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, QV (SEQ ID NO:X), TQV (SEQ ID NO:X), RETQV (SEQ ID NO:X), RRETQV (SEQ ID NO:X), RRRETQV (SEQ ID NO:X), and QRRRETQV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 33 is ETAL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AL (SEQ ID NO:X), TAL (SEQ ID NO:X), RETAL (SEQ ID NO:X), RRETAL (SEQ ID NO:X), GRRETAL (SEQ ID NO:X), and QGRRETAL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 #35 is ETEV (SEQ ID NO:X).

When naturally-occurring residues are added or removed from the core sequence, EV (SEQ ID NO:X), TEV (SEQ ID NO:X), RETEV (SEQ ID NO:X), RRETEV (SEQ ID NO:X), TRRETEV (SEQ ID NO:X), and PTRRETEV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 52 is VTQV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, QV (SEQ ID NO:X), TQV (SEQ ID NO:X), RVTQV (SEQ ID NO:X), RRVTQV (SEQ ID NO:X), GRRVTQV (SEQ ID NO:X), and QGRRVTQV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 #57 is RTSH (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SH (SEQ ID NO:X), TSH (SEQ ID NO:X), LRTSH (SEQ ID NO:X), ALRTSH (SEQ ID NO:X), PALRTSH (SEQ ID NO:X), and APALRTSH (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 58 is QTQV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, QV (SEQ ID NO:X), TQV (SEQ ID NO:X), RQTQV (SEQ ID NO:X), RRQTQV (SEQ ID NO:X), GRRQTQV (SEQ ID NO:X), and QGRRQTQV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 #66 is ESTV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, TV (SEQ ID NO:X), STV (SEQ ID NO:X), TESTV (SEQ ID NO:X), ATESTV (SEQ ID NO:X), QATESTV (SEQ ID NO:X), and RQATESTV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of HPV E6 77 is QSRQ (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RQ (SEQ ID NO:X), SRQ (SEQ ID NO:X), GQSRQ (SEQ ID NO:X), GGQSRQ (SEQ ID NO:X), GGGQSRQ (SEQ ID NO:X), and RGGGQSRQ (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of TAX is ETEV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EV (SEQ ID NO:X), TEV (SEQ ID NO:X), RETEV (SEQ ID NO:X), FRETEV (SEQ ID NO:X),

HFRETEV (SEQ ID NO:X), and KHFRETEV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

9. PL Sequences of Tight Junction Integral Membrane Proteins

A number of tight junction integral membrane proteins contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, Claudin 1, Claudin 2, Claudin 7, Claudin 9, Claudin 10, and Claudin 18.

The C-terminal core sequence of Claudin 1 is KDYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), DYV (SEQ ID NO:X), GKDYV (SEQ ID NO:X), SGKDYV (SEQ ID NO:X), SSGKDYV (SEQ ID NO:X), and PSSGKDYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Claudin 2 is TGYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), GYV (SEQ ID NO:X), LTGYV (SEQ ID NO:X), SLTGYV (SEQ ID NO:X), YSLTGYV (SEQ ID NO:X), and SYSLTGYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Claudin 7 is KEYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), EYV (SEQ ID NO:X), SKEYV (SEQ ID NO:X), SSKEYV (SEQ ID NO:X), NSSKEYV (SEQ ID NO:X), and SNSSKEYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Claudin 9 is RDYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), DYV (SEQ ID NO:X), KRDYV (SEQ ID NO:X), DKRDYV (SEQ ID NO:X), LDKRDYV (SEQ ID NO:X), and GLDKRDYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Claudin 10 is NAYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), AYV (SEQ ID NO:X), KNAYV (SEQ ID NO:X), DKNAYV (SEQ ID NO:X), FDKNAYV (SEQ ID NO:X), and QFDKNAYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Claudin 18 is HDYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), DYV (SEQ ID NO:X), KHDYV (SEQ ID NO:X), SKHDYV (SEQ ID NO:X), PSKHDYV (SEQ ID NO:X), and YPSKHDYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

10. PL Sequences of Cell Adhesion Molecules

A number of cell adhesion molecules contain a PL motif sequence (PL sequence). As used herein, an adhesion protein is a cell surface protein involved in cell-cell interaction by direct contact with cell surface molecules (e.g., transmembrane proteins or surface proteins) on a different cell. Thus, when a cell expressing a PL adhesion protein contacts an appropriate other cell, the PL adhesion protein localizes at the interface of the two cells and directly contacts a cell surface molecule on the second cell. A cell-cell interface is a region where the plasma membranes of two different cells are in close (generally <10 nm, often about 1 nm) apposition. Typically, direct molecular contact means interaction of molecules at distances where Van der Walls forces are significant, generally less than about 1 nm. Inhibition or modulation can occur in a variety of cell types including, endothelial cells, epithelial cells, keratinocytes, hepatocytes and cardiac myocytes.

These molecules include, but are not limited to, Neuroligin, Nectin 2, JAM (junctional adhesion molecule), neurofascin (chicken), and CSPG4 (chondroitin sulfate proteoglycan 4, melanoma-associated).

The C-terminal core sequence of Neuroligin is TTRV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RV (SEQ ID NO:X), TRV (SEQ ID NO:X), STTRV (SEQ ID NO:X), HSTTRV (SEQ ID NO:X), PHSTTRV (SEQ ID NO:X), and LPHSTTRV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Nectin 2 is AMYV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YV (SEQ ID NO:X), MYV (SEQ ID NO:X), RAMYV (SEQ ID NO:X), SRAMYV (SEQ ID NO:X), MSRAMYV (SEQ ID NO:X), and VMSRAMYV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of JAM is SLFV (SEQ ID NO:X). When

naturally-occurring residues are added or removed from the core sequence, FV (SEQ ID NO:X), LFV (SEQ ID NO:X), SSLFV (SEQ ID NO:X), TSSLFV (SEQ ID NO:X), QTSSLFV (SEQ ID NO:X), and KQTSSLFV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of neurofascin is YSLA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LA (SEQ ID NO:X), SLA (SEQ ID NO:X), IYSLA (SEQ ID NO:X), AIYSLA (SEQ ID NO:X), NAIYSLA (SEQ ID NO:X), and VNAIYSLA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CSPG4 is QYWV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, WV (SEQ ID NO:X), YWV (SEQ ID NO:X), GQYWV (SEQ ID NO:X), NGQYWV (SEQ ID NO:X), KNGQYWV (SEQ ID NO:X), and LKNGQYWV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

11. PL Sequences of Neuron Membrane Transport and Organization Molecules

A number of neuron membrane transport and organization molecules contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, Dopamine transporter, noradrenaline transporter, glutamate transporter 3, GABA transporter 3, MINT-1, MINT-2, MINT-3, presenilin-1, and presenilin-2.

The C-terminal core sequence of Dopamine transporter is WLKV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KV (SEQ ID NO:X), LKV (SEQ ID NO:X), HWLKV (SEQ ID NO:X), RHWLKV (SEQ ID NO:X), LRHWLKV (SEQ ID NO:X), and TLRHWLKV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of noradrenaline transporter is WLAI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AI (SEQ ID NO:X), LAI (SEQ ID NO:X), HWLAI (SEQ ID NO:X), QHWLAI (SEQ ID NO:X), LQHWLAI (SEQ ID NO:X), and QLQHWLAI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of glutamate transporter 3 is TSQF (SEQ ID

NO:X). When naturally-occurring residues are added or removed from the core sequence, QF (SEQ ID NO:X), SQF (SEQ ID NO:X), QTSQF (SEQ ID NO:X), TQTSQF (SEQ ID NO:X), FTQTSQF (SEQ ID NO:X), and SFTQTSQF (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GABA transporter 3 is ETHF (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, HF (SEQ ID NO:X), THF (SEQ ID NO:X), KETHF (SEQ ID NO:X), EKETHF (SEQ ID NO:X), TEKETHF (SEQ ID NO:X), and ITEKETHF (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of MINT-1 is PVYI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YI (SEQ ID NO:X), VYI (SEQ ID NO:X), QPVYI (SEQ ID NO:X), EQPVYI (SEQ ID NO:X), QEQPVYI (SEQ ID NO:X), and AQEQPVYI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of MINT-2 is PLYI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YI (SEQ ID NO:X), LYI (SEQ ID NO:X), TPLYI (SEQ ID NO:X), ETPLYI (SEQ ID NO:X), QETPLYI (SEQ ID NO:X), and GQETPLYI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of MINT-3 is PVYL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YL (SEQ ID NO:X), VYL (SEQ ID NO:X), QPVYL (SEQ ID NO:X), EQPVYL (SEQ ID NO:X), QEQPVYL (SEQ ID NO:X), and GQEQPVYL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of presenilin-1 is QFYI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YI (SEQ ID NO:X), FYI (SEQ ID NO:X), HQFYI (SEQ ID NO:X), FHQFYI (SEQ ID NO:X), AFHQFYI (SEQ ID NO:X), and LAFHQFYI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of presenilin-2 is QLYI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, YI (SEQ ID NO:X), LYI (SEQ ID NO:X), HQLYI (SEQ ID NO:X), SHQLYI (SEQ ID NO:X),

ASHQLYI (SEQ ID NO:X), and LASHQLYI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

12. PL Sequences of Receptor Kinases

A number of receptor kinases contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, ephrin A2, ephrin B1, ephrin B2, c-kit receptor, and ErbB-4 receptor.

The C-terminal core sequence of ephrin A2 is GIPI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, PI (SEQ ID NO:X), IPI (SEQ ID NO:X), VGIPI (SEQ ID NO:X), TVGIPI (SEQ ID NO:X), NTVGIPI (SEQ ID NO:X), and VNTVGIPI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ephrin B1 is YYKV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KV (SEQ ID NO:X), YKV (SEQ ID NO:X), IYYKV (SEQ ID NO:X), NIYYKV (SEQ ID NO:X), ANIYYKV (SEQ ID NO:X), and PANIYYKV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ephrin B2 is SVEV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EV (SEQ ID NO:X), VEV (SEQ ID NO:X), QSVEV (SEQ ID NO:X), IQSVEV (SEQ ID NO:X), QIQSVEV (SEQ ID NO:X), and NQIQSVEV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of c-kit receptor is HDDV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DV (SEQ ID NO:X), DDV (SEQ ID NO:X), VHDDV (SEQ ID NO:X), LVHDDV (SEQ ID NO:X), LLVHDDV (SEQ ID NO:X), and PLLVHDDV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ErbB-4 receptor is NTVV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, VV (SEQ ID NO:X), TVV (SEQ ID NO:X), RNTVV (SEQ ID NO:X), HRNTVV (SEQ ID NO:X), RHRNTVV (SEQ ID NO:X), and YRHRNTVV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

13. PL Sequences of Regulators of G-Protein Signaling

A number of regulators of G-protein signaling contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, RGS12 (regulator of G-protein signaling 12), and GAIP (G-alpha interacting protein) RGS 19.

The C-terminal core sequence of RGS12 is ATFV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, FV (SEQ ID NO:X), TFV (SEQ ID NO:X), HATFV (SEQ ID NO:X), HHATFV (SEQ ID NO:X), AHHATFV (SEQ ID NO:X), and SAHHATFV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of GAIP (G-alpha interacting protein) RGS 19 is SSEA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EA (SEQ ID NO:X), SEA (SEQ ID NO:X), QSSEA (SEQ ID NO:X), SQSSEA (SEQ ID NO:X), PSQSSEA (SEQ ID NO:X), and GPSQSSEA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

14. PL Sequences of Ion Channels and Transporters

A number of regulators of ion channels and transporters contain a PL motif sequence (PL sequence). As used herein, an ion channel protein means a transmembrane protein that itself catalyzes the passage of an ion from aqueous solution on one side of a lipid bilayer membrane to aqueous solution on the other side (e.g., by forming a small pore in the membrane). These molecules include, but are not limited to, Kir2.1 (inwardly rect. K+ channel), and Na+/Pi contransporter 2.

The C-terminal core sequence of Kir2.1 is ESEI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EI (SEQ ID NO:X), SEI (SEQ ID NO:X), RESEI (SEQ ID NO:X), RRESEI (SEQ ID NO:X), LRRESEI (SEQ ID NO:X), and PLRRESEI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Na+/Pi contransporter 2 is ATRL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RL (SEQ ID NO:X), TRL (SEQ ID NO:X), NATRL (SEQ ID NO:X), HNATRL (SEQ ID NO:X), HHNATRL (SEQ ID NO:X), and AHHNATRL (SEQ ID NO:X) may also be used

to target a PDZ domain-containing protein in cells.

15. PL Sequences of Tumor Suppressor Proteins, Cell Viability Associated Proteins, Receptors, and Critical Regulators

A number of tumor suppressor proteins, cell viability associated proteins, receptors, and critical regulators contain a PL motif sequence (PL sequence). These molecules include, but are not limited to, alpha-1-syntrophin, ropporin, CX43 (connexin 43), CD68, a-actinin 2, zona occludens 3 (ZO-3), KIA 1481, CFTCR (cystic fibrosis transmembrane conductance regulator), ActRIIA, CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA, RA-GEF (ras/rap1A-assoc.-GEF), PDZ-binding kinase (PBK), RhoGAP (PTPL1-associated), CITRON protein, Nedasin (s-form), APC-adenomatous polyposis coli protein, CKR5 (HIV Co-receptor), catenin –delta 2, bone morphogenetic protein receptor, TRAF2, Glycophorin C, and PTEN.

The C-terminal core sequence of alpha-1-syntrophin is GLLA (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LA (SEQ ID NO:X), LLA (SEQ ID NO:X), LGLLA (SEQ ID NO:X), RLGLLA (SEQ ID NO:X), TRLGLLA (SEQ ID NO:X), and VTRLGLLA (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ropprin is VQLE (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LE (SEQ ID NO:X), QLE (SEQ ID NO:X), RVQLE (SEQ ID NO:X), PRVQLE (SEQ ID NO:X), NPRVQLE (SEQ ID NO:X), and QNPRVQLE (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ropprin is VQLE (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, LE (SEQ ID NO:X), QLE (SEQ ID NO:X), RVQLE (SEQ ID NO:X), PRVQLE (SEQ ID NO:X), NPRVQLE (SEQ ID NO:X), and QNPRVQLE (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CX43 (connexin 43) is DLEI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, EI (SEQ ID NO:X), LEI (SEQ ID NO:X), DDLEI (SEQ ID NO:X), PDDLEI (SEQ ID NO:X), RPDDLEI (SEQ ID NO:X), and PRPDDLEI (SEQ ID NO:X) may also be used to

target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CD68 is YQAL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AL (SEQ ID NO:X), QAL (SEQ ID NO:X), AYQAL (SEQ ID NO:X), SAYQAL (SEQ ID NO:X), PSAYQAL (SEQ ID NO:X), and RPSAYQAL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of a-actinin 2 is ESDL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DL (SEQ ID NO:X), SDL (SEQ ID NO:X), GESDL (SEQ ID NO:X), YGESDL (SEQ ID NO:X), LYGESDL (SEQ ID NO:X), and ALYGESDL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of zona occludens 3 (ZO-3) is ATDL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DL (SEQ ID NO:X), TDL (SEQ ID NO:X), PATDL (SEQ ID NO:X), GPATDL (SEQ ID NO:X), WGPATDL (SEQ ID NO:X), and DWGPATDL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of KIA 1481 is TSPL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, PL (SEQ ID NO:X), SPL (SEQ ID NO:X), PTSPL (SEQ ID NO:X), GPTSPL (SEQ ID NO:X), WGPTSPL (SEQ ID NO:X), and DWGPTSPL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CFTCR (cystic fibrosis transmembrane conductance regulator) is DTRL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, RL (SEQ ID NO:X), TRL (SEQ ID NO:X), QDTRL (SEQ ID NO:X), VQDTRL (SEQ ID NO:X), EVQDTRL (SEQ ID NO:X), and EEVQDTRL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of ActRIIA is ESSL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SL (SEQ ID NO:X), SSL (SEQ ID NO:X), KESSL (SEQ ID NO:X), PKESSL (SEQ ID NO:X), PPKESSL (SEQ ID NO:X), and FPPKESSL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CAPON (carboxy-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA is EIAV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AV (SEQ ID NO:X), IAV (SEQ ID NO:X), DEIAV (SEQ ID NO:X), DDEIAV (SEQ ID NO:X), LDDEIAV (SEQ ID NO:X), and GLDDEIAV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of RA-GEF (ras/rap1A-assoc.-GEF) is VSAV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, AV (SEQ ID NO:X), SAV (SEQ ID NO:X), QVSAV (SEQ ID NO:X), EQVSAV (SEQ ID NO:X), DEQVSAV (SEQ ID NO:X), and EDEQVSAV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of PDZ-binding kinase (PBK) is ETDV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, DV (SEQ ID NO:X), TDV (SEQ ID NO:X), LETDV (SEQ ID NO:X), ALETDV (SEQ ID NO:X), EALETDV (SEQ ID NO:X), and VEALETDV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of RhoGAP 1 (PTPL1-associated) is PQFV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, FV (SEQ ID NO:X), QFV (SEQ ID NO:X), IPQFV (SEQ ID NO:X), EIPQFV (SEQ ID NO:X), DEIPQFV (SEQ ID NO:X), and EDEIPQFV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CITRON protein is QSSV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SV (SEQ ID NO:X), SSV (SEQ ID NO:X), DQSSV (SEQ ID NO:X), WDQSSV (SEQ ID NO:X), VWDQSSV (SEQ ID NO:X), and KVWDQSSV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Nedasin (s-form) is SSSV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SV (SEQ ID NO:X), SSV (SEQ ID NO:X), FSSSV (SEQ ID NO:X), PFSSSV (SEQ ID NO:X), VPFSSSV (SEQ ID NO:X), and VVPFSSSV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of APC- adenomatous polyposis coli protein

is VTSV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, SV (SEQ ID NO:X), TSV (SEQ ID NO:X), LVTSV (SEQ ID NO:X), YLVTSV (SEQ ID NO:X), SYLVTSV (SEQ ID NO:X), and GSYLVTSV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of CKR5 (HIV Co-receptor) is SVGL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, GL (SEQ ID NO:X), VGL (SEQ ID NO:X), ISVGL (SEQ ID NO:X), EISVGL (SEQ ID NO:X), QEISVGL (SEQ ID NO:X), and EQEISVGL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of cantenin – delta 2 is DSWV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, WV (SEQ ID NO:X), SWV (SEQ ID NO:X), PDSWV (SEQ ID NO:X), SPDSWV (SEQ ID NO:X), ASPDSWV (SEQ ID NO:X), and PASPDSWV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of bone morphogenetic protein receptor is DVKI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KI (SEQ ID NO:X), VKI (SEQ ID NO:X), QDVKI (SEQ ID NO:X), SQDVKI (SEQ ID NO:X), ESQDVKI (SEQ ID NO:X), and VESQDVKI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of TRAF2 is LTGL (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, GL (SEQ ID NO:X), TGL (SEQ ID NO:X), DLTGL (SEQ ID NO:X), VDLTGL (SEQ ID NO:X), IVDLTGL (SEQ ID NO:X), and AIVDLTGL (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of Glycophorin C is EYFI (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, FI (SEQ ID NO:X), YFI (SEQ ID NO:X), KEYFI (SEQ ID NO:X), RKEYFI (SEQ ID NO:X), SRKEYFI (SEQ ID NO:X), and SSRKEYFI (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

The C-terminal core sequence of PTEN is ITKV (SEQ ID NO:X). When naturally-occurring residues are added or removed from the core sequence, KV (SEQ ID NO:X), TKV (SEQ ID NO:X), QITKV (SEQ ID NO:X), TQITKV (SEQ ID NO:X),

HTQITKV (SEQ ID NO:X), and QHTQITKV (SEQ ID NO:X) may also be used to target a PDZ domain-containing protein in cells.

16. Others

The PL proteins that have been identified herein as interacting with particular PDZ proteins also include intracellular proteins, and cytokine receptors, and adaptor proteins. As used herein, an intercellular (i.e., cytosolic) protein has the normal meaning in the art and refers to a protein that is not membrane bound, e.g., has no transmembrane domain. The term cytokine receptor as used herein a cytokine receptor has the normal meaning in the art and refers to a membrane protein with an extracellular domain that specifically binds a cytokine. As used herein, an adaptor protein means a molecule (e.g., protein) that contributes to the formation of a multimolecular complex by binding two or more other biomolecules. The binding of the two or more other molecules by the adaptor molecule/protein generally involves direct molecular contact between the adaptor protein and each of the two or more other molecules.

V. Detection of PDZ Domain-Containing Proteins

As noted *supra*, the present inventors have identified a number of PDZ protein and PL protein interactions that can play a role in modulation of a number of biological functions in a variety of cell types. A comprehensive list of PDZ domain-containing proteins was retrieved from the Sanger Centre database (Pfam) searching for the keyword, "PDZ". The corresponding cDNA sequences were retrieved from GenBank using the NCBI "entrez" database (hereinafter, "GenBank PDZ protein cDNA sequences"). The DNA portion encoding PDZ domains was identified by alignment of cDNA and protein sequence using CLUSTALW. Based on the DNA/protein alignment information, primers encompassing the PDZ domains were designed. The expression of certain PDZ-containing proteins in cells was detected by polymerase chain reaction ("PCR") amplification of cDNAs obtained by reverse transcription ("RT") of cell-derived RNA (i.e., "RT-PCR"). PCR, RT-PCR and other methods for analysis and manipulation of nucleic acids are well known and are described generally in Sambrook et al., (1989) MOLECULAR CLONING: A LABORATORY MANUAL, 2ND ED., VOLS. 1-3, Cold Spring Harbor Laboratory hereinafter, "Sambrook"); and Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing and Wiley-Interscience, New York (1997), as

supplemented through January 1999 (hereinafter "Ausubel").

Samples of cDNA for those sequences identified through the foregoing search were obtained and then amplified. In general, a sample of the cDNA (typically, 1/5 of a 20 µl reaction) was used to conduct PCR. PCR was conducted using primers designed specifically to amplify PDZ domain-containing regions of PDZ proteins of interest. Oligonucleotide primers were designed to amplify one or more PDZ-encoding domains. The DNA sequences encoding the various PDZ domains of interest were identified by inspection (i.e., conceptual translation of the PDZ protein cDNA sequences obtained from GenBank, followed by alignment with the PDZ domain amino acid sequence). **TABLE 9** shows the PDZ-encoded domains amplified, and the GenBank accession number of the PDZ-domain containing proteins. To facilitate subsequent cloning of PDZ domains, the PCR primers included endonuclease restriction sequences at their ends to allow ligation with pGEX-3X cloning vector (Pharmacia, GenBank XXI13852) in frame with glutathione-S transferase (GST).

VI. Assays for Detection of Interactions Between PDZ-Domain Polypeptides and Candidate PDZ Ligand proteins (PL proteins)

Two complementary assays, termed "A' and "G,"" were developed to detect binding between a PDZ-domain polypeptide and candidate PDZ ligand. In each of the two different assays, binding is detected between a peptide having a sequence corresponding to the C-terminus of a protein anticipated to bind to one or more PDZ domains (i.e. a candidate PL peptide) and a PDZ-domain polypeptide (typically a fusion protein containing a PDZ domain). In the "A" assay, the candidate PL peptide is immobilized and binding of a soluble PDZ-domain polypeptide to the immobilized peptide is detected (the "A'" assay is named for the fact that in one embodiment an avidin surface is used to immobilize the peptide). In the "G" assay, the PDZ-domain polypeptide is immobilized and binding of a soluble PL peptide is detected (The "G" assay is named for the fact that in one embodiment a GST-binding surface is used to immobilize the PDZ-domain polypeptide). Preferred embodiments of these assays are described in detail *infra*. However, it will be appreciated by ordinarily skilled practitioners that these assays can be modified in numerous ways while remaining useful for the purposes of the present invention.

A. Production of Fusion Proteins Containing PDZ-Domains

GST-PDZ domain fusion proteins were prepared for use in the assays of the invention. PCR products containing PDZ encoding domains (as described *supra*) were subcloned into an expression vector to permit expression of fusion proteins containing a PDZ domain and a heterologous domain (i.e., a glutathione-S transferase sequence, "GST"). PCR products (i.e., DNA fragments) representing PDZ domain encoding DNA was extracted from agarose gels using the "sephaglas" gel extraction system (Pharmacia) according to the manufacturer's recommendations.

As noted *supra*, PCR primers were designed to include endonuclease restriction sites to facilitate ligation of PCR fragments into a GST gene fusion vector (pGEX-3X; Pharmacia, GenBank accession no. XXU13852) in-frame with the glutathione-S transferase coding sequence. This vector contains an IPTG inducible lacZ promoter. The pGEX-3X vector was linearized using *Bam* HI and *Eco* RI or, in some cases, *Eco* RI or *Sma* I, and dephosphorylated. For most cloning approaches, double digestion with Bam HI and Eco RI was performed, so that the ends of the PCR fragments to clone were Bam HI and Eco RI. In some cases, restriction endonuclease combinations used were Bgl II and Eco RI, Bam HI and Mfe I, or Eco RI only, Sma I only, or BamHI only. When more than one PDZ domain was cloned, the DNA portion cloned represents the PDZ domains and the cDNA portion located between individual domains. Precise locations of cloned fragments used in the assays are indicated in **TABLE 9**. DNA linker sequences between the GST portion and the PDZ domain containing DNA portion vary slightly, dependent on which of the above described cloning sites and approaches were used. As a consequence, the amino acid sequence of the GST-PDZ fusion protein varies in the linker region between GST and PDZ domain. Protein linker sequences corresponding to different cloning sites/approaches are shown below. Linker sequences (vector DNA encoded) are bold, PDZ domain containing gene derived sequences are in italics.

1) **GST—BamHI/***BamHI— PDZ domain insert*
Gly–Ile—*PDZ domain insert*

2) **GST—BamHI/***BglII—PDZ domain insert*
Gly—Ile—*PDZ domain insert*

3) **GST—EcoRI/***EcoI—PDZ domain insert*
Gly—Ile—Pro—Gly–Asn—*PDZ domain insert*

4) **GST–SmaI/***SmaI—PDZ domain insert*
Gly—Ile—Pro—*PDZ domain insert*

The PDZ-encoding PCR fragment and linearized pGEX-3X vector were ethanol precipitated and resuspended in 10 ul standard ligation buffer. Ligation was performed for 4-10 hours at 7°C using T4 DNA ligase. It will be understood that some of the resulting constructs include very short linker sequences and that, when multiple PDZ domains were cloned, the constructs included some DNA located between individual PDZ domains.

The ligation products were transformed in DH5α or BL-21 *E.coli* bacteria strains. Colonies were screened for presence and identity of the cloned PDZ domain containing DNA as well as for correct fusion with the glutathione S-transferase encoding DNA portion by PCR and by sequence analysis. Positive clones were tested in a small-scale assay for expression of the GST/PDZ domain fusion protein and, if expressing, these clones were subsequently grown up for large scale preparations of GST/PDZ fusion protein.

GST-PDZ domain fusion protein was overexpressed following addition of IPTG to the culture medium and purified. Detailed procedure of small scale and large-scale fusion protein expression and purification are described in "GST Gene Fusion System" (second edition, revision 2; published by Pharmacia). In brief, a small culture (50mls) containing a bacterial strain (DH5α, BL21 or JM109) with the fusion protein construct was grown overnight in 2xYT media at 37°C with the appropriate antibiotic selection (100ug/ml ampicillin; a.k.a. 2xYT-amp). The overnight culture was poured into a fresh preparation of 2xYT-amp (typically 1 liter) and grown until the optical density (OD) of the culture was between 0.5 and 0.9 (approximately 2.5 hours). IPTG (isopropyl β-D-thiogalactopyranoside) was added to a final concentration of 1.0mM to induce production of GST fusion protein, and culture was grown an additional 1 hour. All following steps, including centrifugation, were performed on ice or at 4°C. Bacteria were collected by centrifugation (4500 g) and resuspended in Buffer A- (50mM Tris, pH 8.0, 50mM dextrose, 1mM EDTA, 200uM phenylmethylsulfonylfluoride). An equal volume of Buffer A+ (Buffer A-, 4mg/ml lysozyme) was added and incubated on ice for 3 min to lyse bacteria, or until lysis had begun. An equal volume of Buffer B (10mM Tris, pH 8.0, 50mM KCl, 1mM EDTA. 0.5% Tween-20, 0.5% NP40 (a.k.a. IGEPAL CA-630), 200uM phenylmethylsulfonylfluoride) was added and incubated for an additional 20 min on ice. The bacterial cell lysate was centrifuged (x20,000g), and supernatant was run over a column containing 20ml Sepharose CL-4B (Pharmacia) "precolumn beads," i.e., sepharose beads without conjugated glutathione that had been previously washed with 3 bed volumes PBS. The flow-through was added to glutathione Sepharose 4B (Pharmacia, cat no. 17-0765-

01) previously swelled (rehydrated) in 1X phosphate-buffered saline (PBS) and incubated while rotating for 30min-1hr. The supernatant-Sepharose slurry was poured into a column and washed with at least 20 bed volumes of 1X PBS. GST fusion protein was eluted off the glutathione sepharose by applying 0.5-1.0 ml aliquots of 5mM glutathione and collected as separate fractions. Concentrations of fractions were determined by reading absorbance at 280nm and calculating concentration using the absorbance and extinction coefficient. Those fractions containing the highest concentration of fusion protein were pooled and an equal volume of 70% glycerol was added to a final concentration of 35% glycerol. Fusion proteins were assayed for size and quality by SDS gel electrophoresis (PAGE) as described in "Sambrook." Fusion protein aliquots were stored at minus 80°C and at minus 20°C.

B. Identification of Candidate PL Proteins and Synthesis of Peptides

Certain PDZ domains are bound by the C-terminal residues of PDZ-binding proteins. To identify PL proteins the C-terminal residues of sequences were visually inspected for sequences that one might predict would bind to PDZ-domain containing proteins (see, e.g., Doyle et al., 1996, *Cell* 85, 1067; Songyang et al., 1997, *Science* 275, 73), including the additional consenses for PLs identified at Arbor Vita Corporation (**TABLE 8**, and data not shown). **TABLE 8** lists some of these proteins, and provides corresponding C-terminal sequences and GenBank accession numbers.

Synthetic peptides of defined sequence (e.g., corresponding to the carboxyl-termini of the indicated proteins) can be synthesized by any standard resin-based method (see, e.g., U. S. Pat. No. 4,108,846; see also, Caruthers et al., 1980, *Nucleic Acids Res. Symp. Ser.*, 215-223; Horn et al., 1980, *Nucleic Acids Res. Symp. Ser.*, 225-232; Roberge, et al., 1995, *Science* 269:202). The peptides used in the assays described herein were prepared by the FMOC (see, e.g., Guy and Fields, 1997, *Meth. Enz.* 289:67-83; Wellings and Atherton, 1997, *Meth. Enz.*289:44-67). In some cases (e.g., for use in the A and G assays of the invention), peptides were labeled with biotin at the amino-terminus by reaction with a four-fold excess of biotin methyl ester in dimethylsulfoxide with a catalytic amount of base. The peptides were cleaved from the resin using a halide containing acid (e.g. trifluoroacetic acid) in the presence of appropriate antioxidants (e.g. ethanedithiol) and excess solvent lyophilized.

Following lyophilization, peptides can be redissolved and purified by reverse phase high performance liquid chromatography (HPLC). One appropriate HPLC solvent

system involves a Vydac C-18 semi-preparative column running at 5 mL per minute with increasing quantities of acetonitrile plus 0.1% trifluoroacetic acid in a base solvent of water plus 0.1% trifluoroacetic acid. After HPLC purification, the identities of the peptides are confirmed by MALDI cation-mode mass spectrometry. As noted, exemplary biotinylated peptides are provided in **TABLE 8.**

C. Detecting PDZ-PL Interactions

The present inventors were able in part to identify the interactions summarized in **TABLE 7** and **TABLE 12** by developing new high throughput screening assays which are described in greater detail infra. Various other assay formats known in the art can be used to select ligands that are specifically reactive with a particular protein. For example, solid-phase ELISA immunoassays, immunoprecipitation, Biacore, and Western blot assays can be used to identify peptides that specifically bind PDZ-domain polypeptides. As discussed *supra*, two different, complementary assays were developed to detect PDZ-PL interactions. In each, one binding partner of a PDZ-PL pair is immobilized, and the ability of the second binding partner to bind is determined. These assays, which are described *infra*, can be readily used to screen for hundreds to thousand of potential PDZ-ligand interactions in a few hours. Thus these assays can be used to identify yet more novel PDZ-PL interactions in hematopoietic cells. In addition, they can be used to identify antagonists of PDZ-PL interactions (see *infra*).

In some assays, fusion proteins are used in the assays and devices of the invention. Methods for constructing and expressing fusion proteins are well known. Fusion proteins generally are described in Ausubel et al., supra, Kroll et al., 1993, DNA Cell. Biol. 12:441, and Imai et al., 1997, *Cell* 91:521-30. Usually, the fusion protein includes a domain to facilitate immobilization of the protein to a solid substrate ("an immobilization domain"). Often, the immobilization domain includes an epitope tag (i.e., a sequence recognized by an antibody, typically a monoclonal antibody) such as polyhistidine (Bush et al, 1991, *J. Biol Chem* 266:13811-14), SEAP (Berger et al, 1988, *Gene* 66:1-10), or M1 and M2 flag (see, e.g, U.S. Pat. Nos. 5,011,912; 4,851,341; 4,703,004; 4,782,137). In an embodiment, the immobilization domain is a GST coding region. It will be recognized that, in addition to the PDZ-domain and the particular residues bound by an immobilized antibody, protein A, or otherwise contacted with the surface, the protein (e.g., fusion protein), will contain additional residues. In some embodiments these are residues naturally associated with the PDZ-domain

(i.e., in a particular PDZ-protein) but they can include residues of synthetic (e.g., poly(alanine)) or heterologous origin (e.g., spacers of, e.g., between 10 and 300 residues).

PDZ domain-containing polypeptide used in these methods are typically made by (1) constructing a vector (e.g., plasmid, phage or phagemid) comprising a polynucleotide sequence encoding the desired polypeptide, (2) introducing the vector into an suitable expression system (e.g., a prokaryotic, insect, mammalian, or cell free expression system), (3) expressing the fusion protein and (4) optionally purifying the fusion protein.

Generally, expression of the protein comprises inserting the coding sequence into an appropriate expression vector (i.e., a vector that contains the necessary elements for the transcription and translation of the inserted coding sequence required for the expression system employed, e.g., control elements including enhancers, promoters, transcription terminators, origins of replication, a suitable initiation codon (e.g., methionine), open reading frame, and translational regulatory signals (e.g., a ribosome binding site, a termination codon and a polyadenylation sequence. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used.

The coding sequence of the fusion protein includes a PDZ domain and an immobilization domain as described elsewhere herein. Polynucleotides encoding the amino acid sequence for each domain can be obtained in a variety of ways known in the art; typically the polynucleotides are obtained by PCR amplification of cloned plasmids, cDNA libraries, and cDNA generated by reverse transcription of RNA, using primers designed based on sequences determined by the practitioner or, more often, publicly available (e.g., through GenBank). The primers include linker regions (e.g., sequences including restriction sites) to facilitate cloning and manipulation in production of the fusion construct. The polynucleotides corresponding to the PDZ and immobilization regions are joined in-frame to produce the fusion protein-encoding sequence.

The fusion proteins can be expressed as secreted proteins (e.g., by including the signal sequence encoding DNA in the fusion gene; see, e.g., Lui et al, 1993, *PNAS USA*, 90:8957-61) or as nonsecreted proteins.

In certain assays, the PDZ-containing proteins are immobilized on a solid surface. The substrate to which the polypeptide is bound can have any of a variety of forms, e.g., a microtiter dish, a test tube, a dipstick, a microcentrifuge tube, a bead, a spinnable disk,

and the like. Suitable materials include glass, plastic (e.g., polyethylene, PVC, polypropylene, polystyrene, and the like), protein, paper, carbohydrate, lipip monolayer or supported lipid bilayer, and other solid supports. Other materials that can be employed include ceramics, metals, metalloids, semiconductive materials, cements and the like.

In other assays, the fusion proteins are organized as an array. The term "array," as used herein, refers to an ordered arrangement of immobilized fusion proteins, in which particular different fusion proteins (i.e., having different PDZ domains) are located at different predetermined sites on the substrate. Because the location of particular fusion proteins on the array is known, binding at that location can be correlated with binding to the PDZ domain situated at that location. Immobilization of fusion proteins on beads (individually or in groups) is another particularly useful approach. In some instances, individual fusion proteins are immobilized on beads. In one embodiment, mixtures of distinguishable beads are used. Distinguishable beads are beads that can be separated from each other on the basis of a property such as size, magnetic property, color (e.g., using FACS) or affinity tag (e.g., a bead coated with protein A can be separated from a bead not coated with protein A by using IgG affinity methods). Binding to particular PDZ domain can be determined; similarly, the effect of test compounds (i.e., agonists and antagonists of binding) can be determined.

Methods for immobilizing proteins are known, and include covalent and non-covalent methods. One suitable immobilization method is antibody-mediated immobilization. According to this method, an antibody specific for the sequence of an "immobilization domain" of the PDZ-domain containing protein is itself immobilized on the substrate (e.g., by adsorption). One advantage of this approach is that a single antibody can be adhered to the substrate and used for immobilization of a number of polypeptides (sharing the same immobilization domain). For example, an immobilization domain consisting of poly-histidine (Bush et al, 1991, *J. Biol Chem* 266:13811-14) can be bound by an anti-histidine monoclonal antibody (R&D Systems, Minneapolis, MN); an immobilization domain consisting of secreted alkaline phosphatase ("SEAP") (Berger et al, 1988, *Gene* 66:1-10) can be bound by anti-SEAP (Sigma Chemical Company, St. Louis, MO); an immobilization domain consisting of a FLAG epitope can be bound by anti-FLAG. Other ligand-antiligand immobilization methods are also suitable (e.g., an immobilization domain consisting of protein A sequences (Harlow and Lane, 1988, Antibodies A Laboratory Manual, Cold Spring Harbor Laboratory; Sigma Chemical Co., St. Louis, MO) can be bound by IgG; and an immobilization domain consisting of streptavidin

can be bound by biotin (Harlow & Lane, *supra*; Sigma Chemical Co., St. Louis, MO). In a preferred embodiment, the immobilization domain is a GST moiety, as described herein.

When antibody-mediated immobilization methods are used, glass and plastic are especially useful substrates. The substrates can be printed with a hydrophobic (e.g., Teflon) mask to form wells. Preprinted glass slides with 3, 10 and 21 wells per 14.5 cm^2 slide "working area" are available from, e.g., SPI Supplies, West Chester, PA; also see U.S. Pat. No. 4,011,350). In certain applications, a large format (12.4 cm x 8.3 cm) glass slide is printed in a 96 well format is used; this format facilitates the use of automated liquid handling equipment and utilization of 96 well format plate readers of various types (fluorescent, colorimetric, scintillation). However, higher densities can be used (e.g., more than 10 or 100 polypeptides per cm^2). See, e.g., MacBeath et al, 2000, *Science* 289:1760-63.

Typically, antibodies are bound to substrates (e.g., glass substrates) by adsorption. Suitable adsorption conditions are well known in the art and include incubation of 0.5-50ug/ml (e.g., 10 ug/ml) mAb in buffer (e.g., PBS, or 50 to 300 mM Tris, MOPS, HEPES, PIPES, acetate buffers, pHs 6.5 to 8, at 4°C) to 37°C and from 1hr to more than 24 hours.

Proteins can be covalently bound or noncovalently attached through nonspecific bonding. If covalent bonding between the fusion protein and the surface is desired, the surface will usually be polyfunctional or be capable of being polyfunctionalized. Functional groups which can be present on the surface and used for linking can include carboxylic acids, aldehydes, amino groups, cyano groups, ethylenic groups, hydroxyl groups, mercapto groups and the like. The manner of linking a wide variety of compounds to various surfaces is well known and is amply illustrated in the literature.

<u>"A Assay" Detection of PDZ-Ligand Binding Using Immobilized PL Peptide.</u>

In this particular assay, biotinylated candidate PL peptides are immobilized on an avidin-coated surface. The binding of PDZ-domain fusion protein to this surface is then measured. In certain assays, the PDZ-domain fusion protein is a GST/PDZ fusion protein and the assay is carried out as follows:

(1) Avidin is bound to a surface, e.g. a protein binding surface. In one embodiment, avidin is bound to a polystyrene 96 well plate (e.g., Nunc Polysorb (cat #475094)

by addition of 100 uL per well of 20 ug/mL of avidin (Pierce) in phosphate buffered saline without calcium and magnesium, pH 7.4 ("PBS", GibcoBRL) at 4°C for 12 hours. The plate is then treated to block nonspecific interactions by addition of 200 uL per well of PBS containing 2 g per 100 mL protease-free bovine serum albumin ("PBS/BSA") for 2 hours at 4°C. The plate is then washed 3 times with PBS by repeatedly adding 200 uL per well of PBS to each well of the, plate and then dumping the contents of the plate into a waste container and tapping the plate gently on a dry surface.

(2) Biotinylated PL peptides (or candidate PL peptides, e.g., see TABLE 8) are immobilized on the surface of wells of the plate by addition of 50 uL per well of 0.4 uM peptide in PBS/BSA for 30 minutes at 4°C. Usually, each different peptide is added to at least eight different wells so that multiple measurements (e.g. duplicates and also measurements using different (GST/PDZ-domain fusion proteins and a GST alone negative control) can be made, and also additional negative control wells are prepared in which no peptide is immobilized. Following immobilization of the PL peptide on the surface, the plate is washed 3 times with PBS.

(3) GST/PDZ-domain fusion protein (prepared as described *supra*) is allowed to react with the surface by addition of 50 uL per well of a solution containing 5 ug/mL GST/PDZ-domain fusion protein in PBS/BSA for 2 hours at 4°C. As a negative control, GST alone (i.e. not a fusion protein) is added to specified wells, generally at least 2 wells (i.e. duplicate measurements) for each immobilized peptide. After the 2 hour reaction, the plate is washed 3 times with PBS to remove unbound fusion protein.

(4) The binding of the GST/PDZ-domain fusion protein to the avidin-biotinylated peptide surface can be detected using a variety of methods, and detectors known in the art. In one assay format, 50 uL per well of an anti-GST antibody in PBS/BSA (e.g. 2.5 ug/mL of polyclonal goat-anti-GST antibody, Pierce) is added to the plate and allowed to react for 20 minutes at 4°C. The plate is washed 3 times with PBS and a second, detectably labeled antibody is added. In another assay, 50 uL per well of 2.5 ug/mL of horseradish peroxidase (HRP)-conjugated polyclonal rabbit anti-goat immunoglobulin antibody is added to the plate and allowed to react for 20 minutes at 4°C. The plate is washed 5 times with 50 mM Tris pH

8.0 containing 0.2% Tween 20, and developed by addition of 100 uL per well of HRP-substrate solution (TMB, Dako) for 20 minutes at room temperature (RT). The reaction of the HRP and its substrate is terminated by the addition of 100 uL per well of 1M sulfuric acid and the optical density (O.D.) of each well of the plate is read at 450 nm.

(5) Specific binding of a PL peptide and a PDZ-domain polypeptide is detected by comparing the signal from the well(s) in which the PL peptide and PDZ domain polypeptide are combined with the background signal(s). The background signal is the signal found in the negative controls. Typically a specific or selective reaction will be at least twice background signal, more typically more than 5 times background, and most typically 10 or more times the background signal. In addition, a statistically significant reaction will involve multiple measurements of the reaction with the signal and the background differing by at least two standard errors, more typically four standard errors, and most typically six or more standard errors. Correspondingly, a statistical test (e.g. a T-test) comparing repeated measurements of the signal with repeated measurements of the background will result in a p-value < 0.05, more typically a p-value < 0.01, and most typically a p-value < 0.001 or less.

As noted, in an embodiment of the "A" assay, the signal from binding of a GST/PDZ-domain fusion protein to an avidin surface not exposed to (i.e. not covered with) the PL peptide is one suitable negative control (sometimes referred to as "B"). The signal from binding of GST polypeptide alone (i.e. not a fusion protein) to an avidin-coated surface that has been exposed to (i.e. covered with) the PL peptide is a second suitable negative control (sometimes referred to as "B2"). Because all measurements are done in multiples (i.e. at least duplicate) the arithmetic mean (or, equivalently, average) of several measurements is used in determining the binding, and the standard error of the mean is used in determining the probable error in the measurement of the binding. The standard error of the mean of N measurements equals the square root of the following: the sum of the squares of the difference between each measurement and the mean, divided by the product of (N) and (N-1). Thus, in some assays, specific binding of the PDZ protein to the plate-bound PL peptide is determined by comparing the mean signal ("mean S") and standard error of the signal ("SE") for a particular PL-PDZ combination with the mean B1 and/or mean B2.

<u>"G Assay" - Detection of PDZ-Ligand Binding Using Immobilized PDZ-</u>

Domain Fusion Polypeptide

In other assays, a GST/PDZ fusion protein is immobilized on a surface ("G" assay). The binding of labeled PL peptide (e.g., as listed in **TABLE 8**) to this surface is then measured. Typically, the assay is carried out as follows:

(1) A PDZ-domain polypeptide is bound to a surface, e.g. a protein binding surface. In a preferred embodiment, a GST/PDZ fusion protein containing one or more PDZ domains is bound to a polystyrene 96-well plate. The GST/PDZ fusion protein can be bound to the plate by any of a variety of standard methods known to one of skill in the art, although some care must be taken that the process of binding the fusion protein to the plate does not alter the ligand-binding properties of the PDZ domain. In some instances, the GST/PDZ fusion protein is bound via an anti-GST antibody that is coated onto the 96-well plate. Adequate binding to the plate can be achieved when:

a. 100 uL per well of 5 ug/mL goat anti-GST polyclonal antibody (Pierce) in PBS is added to a polystyrene 96-well plate (e.g., Nunc Polysorb) at 4°C for 12 hours.

b. The plate is blocked by addition of 200 uL per well of PBS/BSA for 2 hours at 4°C.

c. The plate is washed 3 times with PBS.

d. 50 uL per well of 5 ug/mL GST/PDZ fusion protein) or, as a negative control, GST polypeptide alone (i.e. not a fusion protein) in PBS/BSA is added to the plate for 2 hours at 4°C.

e. The plate is again washed 3 times with PBS.

(2) Biotinylated PL peptides are allowed to react with the surface by addition of 50 uL per well of 20 uM solution of the biotinylated peptide in PBS/BSA for 10 minutes at 4°C, followed by an additional 20 minute incubation at 25°C. The plate is washed 3 times with ice cold PBS.

(3) The binding of the biotinylated peptide to the GST/PDZ fusion protein surface can be detected using a variety of methods and detectors known to one of skill in the art. In some assays, 100 uL per well of 0.5 ug/mL streptavidin-horse radish peroxidase (HRP)

conjugate dissolved in BSA/PBS is added and allowed to react for 20 minutes at 4°C. The plate is then washed 5 times with 50 mM Tris pH 8.0 containing 0.2% Tween 20, and developed by addition of 100 uL per well of HRP-substrate solution (TMB, Dako) for 20 minutes at room temperature (RT). The reaction of the HRP and its substrate is terminated by addition of 100 uL per well of 1M sulfuric acid, and the absorbance of each well of the plate is read at 450nm.

(4) Specific binding of a PL peptide and a PDZ domain polypeptide is determined by comparing the signal from the well(s) in which the PL peptide and PDZ domain polypeptide are combined, with the background signal(s). The background signal is the signal found in the negative control(s). Typically a specific or selective reaction will be at least twice background signal, more typically more than 5 times background, and most typically 10 or more times the background signal. In addition, a statistically significant reaction will involve multiple measurements of the reaction with the signal and the background differing by at least two standard errors, more typically four standard errors, and most typically six or more standard errors. Correspondingly, a statistical test (e.g. a T-test) comparing repeated measurements of the signal with -repeated measurements of the background will result in a p-value < 0.05, more typically a p-value < 0.01, and most typically a p-value < 0.001 or less. As noted, in an embodiment of the "G" assay, the signal from binding of a given PL peptide to immobilized (surface bound) GST polypeptide alone is one suitable negative control (sometimes referred to as "B 1"). Because all measurement are done in multiples (i.e. at least duplicate) the arithmetic mean (or, equivalently, average.) of several measurements is used in determining the binding, and the standard error of the mean is used in determining the probable error in the measurement of the binding. The standard error of the mean of N measurements equals the square root of the following: the sum of the squares of the difference between each measurement and the mean, divided by the product of (N) and (N-1). Thus, in some instances, specific binding of the PDZ protein to the platebound peptide is determined by comparing the mean signal ("mean S") and standard error of the signal ("SE") for a particular PL-PDZ combination with the mean B1.

"G' assay" and "G'' assay"

Two specific modifications of the specific conditions described *supra* for the "G assay" can be utilized. The modified assays use lesser quantities of labeled PL peptide and

have slightly different biochemical requirements for detection of PDZ-ligand binding compared to the specific assay conditions described *supra*.

For convenience, the assay conditions described in this section are referred to as the "G' assay" and the "G'' assay," with the specific conditions described in the preceding section on G assays being referred to as the "G^0 assay." The "G' assay" is identical to the "G^0 assay" except at step (2) the peptide concentration is 10 uM instead of 20 uM. This results in slightly lower sensitivity for detection of interactions with low affinity and/or rapid dissociation rate. Correspondingly, it slightly increases the certainty that detected interactions are of sufficient affinity and half-life to be of biological importance and useful therapeutic targets.

The "G'' assay" is identical to the "G^0 assay" except that at step (2) the peptide concentration is 1 uM instead of 20 uM and the incubation is performed for 60 minutes at 25°C (rather than, e.g., 10 minutes at 4°C followed by 20 minutes at 25°C). This results in lower sensitivity for interactions of low affinity, rapid dissociation rate, and/or affinity that is less at 25°C than at 4°C. Interactions will have lower affinity at 25°C than at 4°C if (as we have found to be generally true for PDZ-ligand binding) the reaction entropy is negative (i.e. the entropy of the products is less than the entropy of the reactants). In contrast, the PDZ-PL binding signal can be similar in the "G'' assay" and the "G^0 assay" for interactions of slow association and dissociation rate, as the PDZ-PL complex will accumulate during the longer incubation of the "G'' assay." Thus comparison of results of the "G'' assay" and the "G^0 assay" can be used to estimate the relative entropies, enthalpies, and kinetics of different PDZ-PL interactions. (Entropies and enthalpies are related to binding affinity by the equations delta G = RT ln (Kd) = delta H – T delta S where delta G, H, and S are the reaction free energy, enthalpy, and entropy respectively, T is the temperature in degrees Kelvin, R is the gas constant, and Kd is the equilibrium dissociation constant). In particular, interactions that are detected only or much more strongly in the "G^0 assay" generally have a rapid dissociation rate at 25°C (t1/2 < 10 minutes) and a negative reaction entropy, while interactions that are detected similarly strongly in the "G'' assay" generally have a slower dissociation rate at 25°C (t1/2 > 10 minutes). Rough estimation of the thermodynamics and kinetics of PDZ-PL interactions (as can be achieved via comparison of results of the "G^0 assay" versus the "G'' assay" as outlined *supra)* can be used in the design of efficient inhibitors of the interactions. For example, a small molecule inhibitor based on the chemical structure of a PL that

dissociates slowly from a given PDZ domain (as evidenced by similar binding in the "G'' assay" as in the "G^0 assay") can itself dissociate slowly and thus be of high affinity.

In this manner, variation of the temperature and duration of step (2) of the "G assay" can be used to provide insight into the kinetics and thermodynamics of the PDZ-ligand binding reaction and into design of inhibitors of the reaction.

Assay Variations

As discussed *supra*, it will be appreciated that many of the steps in the above-described assays can be varied, for example, various substrates can be used for binding the PL and PDZ-containing proteins; different types of PDZ containing fusion proteins can be used; different labels for detecting PDZ/PL interactions can be employed; and different ways of detection can be used.

The PDZ-PL detection assays can employ a variety of surfaces to bind the PL and PDZ-containing proteins. For example, a surface can be an "assay plate" which is formed from a material (e.g. polystyrene) which optimizes adherence of either the PL protein or PDZ-containing protein thereto. Generally, the individual wells of the assay plate will have a high surface area to volume ratio and therefore a suitable shape is a flat bottom well (where the proteins of the assays are adherent). Other surfaces include, but are not limited to, polystyrene or glass beads, polystyrene or glass slides, and the like.

For example, the assay plate can be a "microtiter" plate. The term "microtiter" plate when used herein refers to a multiwell assay plate, e.g., having between about 30 to 200 individual wells, usually 96 wells. Alternatively, high-density arrays can be used. Often, the individual wells of the microtiter plate will hold a maximum volume of about 250 ul. Conveniently, the assay plate is a 96 well polystyrene plate (such as that sold by Becton Dickinson Labware, Lincoln Park, N.J.), which allows for automation and high throughput screening. Other surfaces include polystyrene microtiter ELISA plates such as that sold by Nunc Maxisorp, Inter Med, Denmark. Often, about 50 ul to 300 ul, more preferably 100 ul to 200 ul, of an aqueous sample comprising buffers suspended therein will be added to each well of the assay plate.

The detectable labels of the invention can be any detectable compound or composition which is conjugated directly or indirectly with a molecule (such as described above). The label can be detectable by itself (e.g., radioisotope labels or fluorescent labels) or,

in the case of an enzymatic label, can catalyze a chemical alteration of a substrate compound or composition which is detectable. The preferred label is an enzymatic one which catalyzes a color change of a non-radioactive color reagent.

Sometimes, the label is indirectly conjugated with the antibody. One of skill is aware of various techniques for indirect conjugation. For example, the antibody can be conjugated with biotin and any of the categories of labels mentioned above can be conjugated with avidin, or vice versa (see also "A" and "G" assay above). Biotin binds selectively to avidin and thus, the label can be conjugated with the antibody in this indirect manner. See, Ausubel, *supra*, for a review of techniques involving biotin-avidin conjugation and similar assays. Alternatively, to achieve indirect conjugation of the label with the antibody, the antibody is conjugated with a small hapten (e.g. digoxin) and one of the different types of labels mentioned above is conjugated with an anti-hapten antibody (e.g. anti-digoxin antibody). Thus, indirect conjugation of the label with the antibody can be achieved.

Assay variations can include different washing steps. By "washing" is meant exposing the solid phase to an aqueous solution (usually a buffer or cell culture media) in such a way that unbound material (e.g., non-adhering cells, non-adhering capture agent, unbound ligand, receptor, receptor construct, cell lysate, or HRP antibody) is removed therefrom. To reduce background noise, it is convenient to include a detergent (e.g., Triton X) in the washing solution. Usually, the aqueous washing solution is decanted from the wells of the assay plate following washing. Conveniently, washing can be achieved using an automated washing device. Sometimes, several washing steps (e.g., between about 1 to 10 washing steps) can be required.

Various buffers can also be used in PDZ-PL detection assays. For example, various blocking buffers can be used to reduce assay background. The term "blocking buffer" refers to an aqueous, pH buffered solution containing at least one blocking compound which is able to bind to exposed surfaces of the substrate which are not coated with a PL or PDZ-containing protein. The blocking compound is normally a protein such as bovine serum albumin (BSA), gelatin, casein or milk powder and does not cross-react with any of the reagents in the assay. The block buffer is generally provided at a pH between about 7 to 7.5 and suitable buffering agents include phosphate and TRIS.

Various enzyme-substrate combinations can also be utilized in detecting PDZ-PL interactions. Examples of enzyme-substrate combinations include, for example:

(i) Horseradish peroxidase (HRPO) with hydrogen peroxidase as a substrate, wherein the hydrogen peroxidase oxidizes a dye precursor (e.g. orthophenylene diamine [OPD] or 3,3',5,5'-tetramethyl benzidine hydrochloride [TMB]) (as described above).

(ii) alkaline phosphatase (AP) with para-Nitrophenyl phosphate as chromogenic substrate.

(iii) β-D-galactosidase (β D-Gal) with a chromogenic substrate (e.g. p-nitrophenyl- β-D-galactosidase) or fluorogenic substrate 4-methylumbelliferyl- β-D-galactosidase.

Numerous other enzyme-substrate combinations are available to those skilled in the art. For a general review of these, see U.S. Pat. Nos. 4,275,149 and 4,318,980, both of which are herein incorporated by reference.

Further, it will be appreciated that, although, for convenience, the present discussion primarily refers antagonists of PDZ-PL interactions, agonists of PDZ-PL interactions can be identified using the methods disclosed herein or readily apparent variations thereof.

VII. Results of PDZ-PL Interaction Assays

TABLE 7 and **TABLE 12**, *supra*, shows the results of assays in which specific binding was detected using the "G'" assay described herein.

VIII. Measurement of PDZ-Ligand Binding Affinity

The "A" and "G" assays described supra can be used to determine the "apparent affinity" of binding of a PDZ ligand peptide to a PDZ-domain polypeptide. Apparent affinity is determined based on the concentration of one molecule required to saturate the binding of a second molecule (e.g., the binding of a ligand to a receptor). Two particularly useful approaches for quantitation of apparent affinity of PDZ-ligand binding are provided *infra*.

(1) A GST/PDZ fusion protein, as well as GST alone as a negative control, are bound to a surface (e.g., a 96-well plate) and the surface blocked and washed as described supra for the "G" assay.

(2) 50 uL per well of a solution of biotinylated PL peptide (e.g. as shown in **TABLE 8**) is added to the surface in increasing concentrations in PBS/BSA (e.g. at 0.1 uM, 0.33 uM, 1 uM, 3.3 uM, 10 uM, 33 uM, and 100 uM). In some instances, the PL peptide is allowed to react with the bound GST/PDZ fusion protein (as well as the GST alone negative

control) for 10 minutes at 4°C followed by 20 minutes at 25°C. The plate is washed 3 times with ice cold PBS to remove unbound labeled peptide.

(3) The binding of the PL peptide to the immobilized PDZ-domain polypeptide is detected as described supra for the "G" assay.

(4) For each concentration of peptide, the net binding signal is determined by subtracting the binding of the peptide to GST alone from the binding of the peptide to the GST/PDZ fusion protein. The net binding signal is then plotted as a function of ligand concentration and the plot is fit (e.g. by using the Kaleidagraph software package curve fitting algorithm; Synergy Software) to the following equation, where "$\text{Signal}_{[\text{ligand}]}$" is the net binding signal at PL peptide concentration "[ligand]," "Kd" is the apparent affinity of the binding event, and "Saturation Binding" is a constant determined by the curve fitting algorithm to optimize the fit to the experimental data:

$$\text{Signal}_{[\text{ligand}]} = \text{Saturation Binding x ([ligand] / ([ligand] + Kd))}$$

For reliable application of the above equation, it is necessary that the highest peptide ligand concentration successfully tested experimentally be greater than, or at least similar to, the calculated Kd (equivalently, the maximum observed binding should be similar to the calculated saturation binding). In cases where satisfying the above criteria proves difficult, an alternative approach (infra) can be used.

Approach 2:

(1) A fixed concentration of a PDZ-domain polypeptide and increasing concentrations of a labeled PL peptide (labeled with, for example, biotin or fluorescein, see **TABLE 9** for representative peptide amino acid sequences) are mixed together in solution and allowed to react. In certain assays, peptide concentrations are 0.1 uM, 1 uM, 10 uM, 100 uM, 1 mM. In other assays, appropriate reaction times can range from 10 minutes to 2 days at temperatures ranging from 4°C to 37°C. In some instances, the identical reaction can also be carried out using a non-PDZ domain-containing protein as a control (e.g., if the PDZ-domain polypeptide is fusion protein, the fusion partner can be used).

(2) PDZ-ligand complexes can be separated from unbound labeled peptide using a variety of methods known in the art. For example, the complexes can be separated using high performance size-exclusion chromatography (HPSEC, gel filtration) (Rabinowitz

et al., 1998, *Immunity* 9:699), affinity chromatography (e.g., using glutathione Sepharose beads), and affinity absorption (e.g., by binding to an anti-GST-coated plate as described *supra*).

(3) The PDZ-ligand complex is detected based on presence of the label on the peptide ligand using a variety of methods and detectors known to one of skill in the art. For example, if the label is fluorescein and the separation is achieved using HPSEC, an in-line fluorescence detector can be used. The binding can also be detected as described supra for the G assay.

(4) The PDZ-ligand binding signal is plotted as a function of ligand concentration and the plot is fit. (e.g., by using the Kaleidagraph software package curve fitting algorithm) to the following equation, where "$Signal_{[ligand]}$" is the binding signal at PL peptide concentration "[ligand]," "Kd" is the apparent affinity of the binding event, and "Saturation Binding" is a constant determined by the curve fitting algorithm to optimize the fit to the experimental data:

$$Signal_{[Ligand]} = \text{Saturation Binding} \times ([ligand] / ([ligand + Kd])$$

Measurement of the affinity of a labeled peptide ligand binding to a PDZ-domain polypeptide is useful because knowledge of the affinity (or apparent affinity) of this interaction allows rational design of inhibitors of the interaction with known potency. The potency of inhibitors in inhibition would be similar to (i.e., within one-order of magnitude of) the apparent affinity of the labeled peptide ligand binding to the PDZ-domain.

Thus, one method of determining the apparent affinity of binding between a PDZ domain and a ligand involves immobilizing a polypeptide comprising the PDZ domain and a non-PDZ domain on a surface, contacting the immobilized polypeptide with a plurality of different concentrations of the ligand, determining the amount of binding of the ligand to the immobilized polypeptide at each of the concentrations of ligand, and calculating the apparent affinity of the binding based on that data. Typically, the polypeptide comprising the PDZ domain and a non-PDZ domain is a fusion protein. In some instances, the e.g., fusion protein is GST-PDZ fusion protein, but other polypeptides can also be used (e.g., a fusion protein including a PDZ domain and any of a variety of epitope tags, biotinylation signals and the like), so long as the polypeptide can be immobilized in an orientation that does not abolish the ligand binding properties of the PDZ domain, e.g., by tethering the polypeptide to the surface via the

non-PDZ domain via an anti-domain antibody and leaving the PDZ domain as the free end. It was discovered, for example, reacting a PDZ-GST fusion polypeptide directly to a plastic plate provided suboptimal results. The calculation of binding affinity itself can be determined using any suitable equation (e.g., as shown *supra*; also see Cantor and Schimmel (1980) BIOPHYSICAL CHEMISTRY WH Freeman & Co., San Francisco) or software.

Thus, in certain methods, the polypeptide is immobilized by binding the polypeptide to an immobilized immunoglobulin that binds the non-PDZ domain (e.g., an anti-GST antibody when a GST-PDZ fusion polypeptide is used). In some instances, the step of contacting the ligand and PDZ-domain polypeptide is carried out under the conditions provided *supra* in the description of the "G" assay. It will be appreciated that binding assays are conveniently carried out in multiwell plates (e.g., 24-well, 96-well plates, or 384 well plates).

The present method has considerable advantages over other methods for measuring binding affinities PDZ-PL affinities, which typically involve contacting varying concentrations of a GST-PDZ fusion protein to a ligand–coated surface. For example, some previously described methods for determining affinity (e.g., using immobilized ligand and GST-PDZ protein in solution) did not account for oligomerization state of the fusion proteins used, resulting in potential errors of more than an order of magnitude.

Although not sufficient for quantitative measurement of PDZ-PL binding affinity, an estimate of the relative strength of binding of different PDZ-PL pairs can be made based on the absolute magnitude of the signals observed in the "G assay." This estimate reflects several factors, including biologically relevant aspects of the interaction, including the affinity and the dissociation rate. For comparisons of different ligands binding to a given PDZ domain-containing protein, differences in absolute binding signal likely relate primarily to the affinity and/or dissociation rate of the interactions of interest.

IX. Assays to Identify Novel PDZ Domain Binding Moieties and to Identify Modulator of PDZ Protein-PL Protein Binding

Although described *supra* primarily in terms of identifying interactions between PDZ-domain polypeptides and PL proteins, the assays described *supra* and other assays can also be used to identify the binding of other molecules (e.g., peptide mimetics, small molecules, and the like) to PDZ domain sequences. For example, using the assays disclosed herein, combinatorial and other libraries of compounds can be screened, e.g., for molecules that specifically bind to PDZ domains. Screening of libraries can be accomplished by any of a

variety of commonly known methods. See, *e.g.*, the following references, which disclose screening of peptide libraries: Parmley and Smith, 1989, *Adv. Exp. Med. Biol.* 251:215-218; Scott and Smith, 1990, *Science* 249:386-390; Fowlkes et al., 1992; *BioTechniques* 13:422-427; Oldenburg et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:5393-5397; Yu et al., 1994, *Cell* 76:933-945; Staudt et al., 1988, *Science* 241:577-580; Bock et al., 1992, *Nature* 355:564-566; Tuerk et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:6988-6992; Ellington et al., 1992, *Nature* 355:850-852; U.S. Patent No. 5,096,815, U.S. Patent No. 5,223,409, and U.S. Patent No. 5,198,346, all to Ladner et al.; Rebar and Pabo, 1993, *Science* 263:671-673; and PCT Publication No. WO 94/18318.

In certain assays, screening can be carried out by contacting the library members with a PDZ-domain polypeptide immobilized on a solid support (e.g. as described *supra* in the "G" assay) and harvesting those library members that bind to the protein. Examples of such screening methods, termed "panning" techniques are described by way of example in Parmley and Smith, 1988, *Gene* 73:305-318; Fowlkes et al., 1992, *BioTechniques* 13:422-427; PCT Publication No. WO 94/18318; and in references cited hereinabove.

In other assays, the two-hybrid system for selecting interacting proteins in yeast (Fields and Song, 1989, *Nature* 340:245-246; Chien et al., 1991, *Proc. Natl. Acad. Sci. USA* 88:9578-9582) is used to identify molecules that specifically bind to a PDZ domain-containing protein. Furthermore, the identified molecules are further tested for their ability to inhibit transmembrane receptor interactions with a PDZ domain.

In one aspect of the invention, antagonists of an interaction between a PDZ protein and a PL protein are identified. In one embodiment, a modification of the "A" assay described *supra* is used to identify antagonists. In one embodiment, a modification of the "G" assay described *supra* is used to identify antagonists.

Screening assays such as these can be used to detect molecules that specifically bind to PDZ domains. Such molecules are useful as agonists or antagonists of PDZ-protein-mediated cell function (e.g., cell activation, e.g., T cell activation, vesicle transport, cytokine release, growth factors, transcriptional changes, cytoskeleton rearrangement, cell movement, chemotaxis, and the like). Thus assays to detect molecules that specifically bind to PDZ domain-containing proteins are provided. For example, recombinant cells expressing PDZ domain-encoding nucleic acids can be used to produce PDZ domains in these assays and to screen for molecules that bind to the domains. Molecules are contacted with the PDZ domain

(or fragment thereof) under conditions conducive to binding, and then molecules that specifically bind to such domains are identified. Methods that can be used to carry out the foregoing are commonly known in the art.

It will be appreciated by the ordinarily skilled practitioner that, in some assays, antagonists are identified by conducting the A or G assays in the presence and absence of a known or candidate antagonist. When decreased binding is observed in the presence of a compound, that compound is identified as an antagonist. Increased binding in the presence of a compound signifies that the compound is an agonist.

For example, in one assay, a test compound can be identified as an inhibitor (antagonist) of binding between a PDZ protein and a PL protein by contacting a PDZ domain polypeptide and a PL peptide in the presence and absence of the test compound, under conditions in which they would (but for the presence of the test compound) form a complex, and detecting the formation of the complex in the presence and absence of the test compound. It will be appreciated that less complex formation in the presence of the test compound than in the absence of the compound indicates that the test compound is an inhibitor of a PDZ protein -PL protein binding.

In certain assays, the "G" assay is used in the presence or absence of a candidate inhibitor. In one embodiment, the "A" assay is used in the presence or absence of a candidate inhibitor.

In other assays (in which a G assay is used), one or more PDZ domain-containing GST-fusion proteins are bound to the surface of wells of a 96-well plate as described *supra* (with appropriate controls including nonfusion GST protein). All fusion proteins are bound in multiple wells so that appropriate controls and statistical analysis can be done. A test compound in BSA/PBS (typically at multiple different concentrations) is added to wells. Immediately thereafter, 30 uL of a detectably labeled (e.g., biotinylated) peptide known to bind to the relevant PDZ domain (see, e.g., **TABLE 7** and **TABLE 12**) is added in each of the wells at a final concentration of, e.g., between about 2 uM and about 40 uM, typically 5 uM, 15 uM, or 25 uM. This mixture is then allowed to react with the PDZ fusion protein bound to the surface for 10 minutes at 4°C followed by 20 minutes at 25°C. The surface is washed free of unbound peptide three times with ice cold PBS and the amount of binding of the peptide in the presence and absence of the test compound is determined. Usually, the level of binding is measured for each set of replica wells (e.g. duplicates) by subtracting the mean GST alone

background from the mean of the raw measurement of peptide binding in these wells.

In certain assays, the A assay is carried out in the presence or absence of a test candidate to identify inhibitors of PL-PDZ interactions.

In some approaches, a test compound is determined to be a specific inhibitor of the binding of the PDZ domain (P) and a PL (L) sequence when, at a test compound concentration of less than or equal to 1 mM (e.g., less than or equal to: 500 uM, 100 uM, 10 uM, 1 uM, 100 nM or 1 nM), the binding of P to L in the presence of the test compound is less than about 50% of the binding in the absence of the test compound (in various embodiments, less than about 25%, less than about 10%, or less than about 1%). Preferably, the net signal of binding of P to L in the presence of the test compound plus six (6) times the standard error of the signal in the presence of the test compound is less than the binding signal in the absence of the test compound.

In one approach, assays for an inhibitor are carried out using a single PDZ protein-PL protein pair (e.g., a PDZ domain fusion protein and a PL peptide). In a related approach, the assays are carried out using a plurality of pairs, such as a plurality of different pairs listed in **TABLE 7 or TABLE 12.**

In some instances, it is desirable to identify compounds that, at a given concentration, inhibit the binding of one PL-PDZ pair, but do not inhibit (or inhibit to a lesser degree) the binding of a specified second PL-PDZ pair. These antagonists can be identified by carrying out a series of assays using a candidate inhibitor and different PL-PDZ pairs (e.g., as shown in the matrix of **TABLE 7 or TABLE 12**) and comparing the results of the assays. All such pairwise combinations are contemplated (e.g., test compound inhibits binding of PL_1 to PDZ_1 to a greater degree than it inhibits binding of PL_1 to PDZ_2 or PL_2 to PDZ_2). Importantly, it will be appreciated that, based on the data provided in **TABLE 7** and **TABLE 12** and disclosed elsehwere herein (and additional data that can be generated using the methods described herein) inhibitors with different specificities can readily be designed.

For example, the Ki ("potency") of an inhibitor of a PDZ-PL interaction can be determined. Ki is a measure of the concentration of an inhibitor required to have a biological effect. For example, administration of an inhibitor of a PDZ-PL interaction in an amount sufficient to result in an intracellular inhibitor concentration of at least between about 1 and about 100 Ki is expected to inhibit the biological response mediated by the target PDZ-PL

interaction. The Kd measurement of PDZ-PL binding as determined using the methods *supra* can be used in determining Ki.

Thus, certain methods of determining the potency (Ki) of an inhibitor or suspected inhibitor of binding between a PDZ domain and a ligand involve immobilizing a polypeptide comprising the PDZ domain and a non-PDZ domain on a surface, contacting the immobilized polypeptide with a plurality of different mixtures of the ligand and inhibitor, wherein the different mixtures comprise a fixed amount of ligand and different concentrations of the inhibitor, determining the amount of ligand bound at the different concentrations of inhibitor, and calculating the Ki of the binding based on the amount of ligand bound in the presence of different concentrations of the inhibitor. In some instances, the polypeptide is immobilized by binding the polypeptide to an immobilized immunoglobulin that binds the non-PDZ domain. This method, which is based on the "G" assay described *supra*, is particularly suited for high-throughput analysis of the Ki for inhibitors of PDZ-ligand interactions. Further, using this method, the inhibition of the PDZ-ligand interaction itself is measured, without distortion of measurements by avidity effects.

Typically, at least a portion of the ligand is detectably labeled to permit easy quantitation of ligand binding.

It will be appreciated that the concentration of ligand and concentrations of inhibitor are selected to allow meaningful detection of inhibition. Thus, the concentration of the ligand whose binding is to be blocked is close to or less than its binding affinity (e.g., in other instances less than the 5x Kd of the interaction, in other instances less than 2x Kd, and in still other instances less than 1x Kd). Thus, the ligand is typically present at a concentration of less than 2 Kd (e.g., between about 0.01 Kd and about 2 Kd) and the concentrations of the test inhibitor typically range from 1 nM to 100 uM (e.g. a 4-fold dilution series with highest concentration 10 uM or 1 mM). In a preferred embodiment, the Kd is determined using the assay disclosed *supra*.

The Ki of the binding can be calculated by any of a variety of methods routinely used in the art, based on the amount of ligand bound in the presence of different concentrations of the inhibitor. In an illustrative embodiment, for example, a plot of labeled ligand binding versus inhibitor concentration is fit to the equation:

$$S_{inhibitor} = S_0 * Ki/([I]+Ki)$$

where $S_{inhibitor}$ is the signal of labeled ligand binding to immobilized PDZ domain in the presence of inhibitor at concentration [I] and S_0 is the signal in the absence of inhibitor (i.e., [I] = 0). Typically [I] is expressed as a molar concentration.

In certain methods, an enhancer (sometimes referred to as, augmentor or agonist) of binding between a PDZ domain and a ligand is identified by immobilizing a polypeptide comprising the PDZ domain and a non-PDZ domain on a surface, contacting the immobilized polypeptide with the ligand in the presence of a test agent and determining the amount of ligand bound, and comparing the amount of ligand bound in the presence of the test agent with the amount of ligand bound by the polypeptide in the absence of the test agent. At least two-fold (often at least 5-fold) greater binding in the presence of the test agent compared to the absence of the test agent indicates that the test agent is an agent that enhances the binding of the PDZ domain to the ligand. As noted *supra*, agents that enhance PDZ-ligand interactions are useful for disruption (dysregulation) of biological events requiring normal PDZ-ligand function (e.g., cancer cell division and metastasis, and activation and migration of immune cells).

The "potency" or "$K_{enhancer}$" of an enhancer of a PDZ- ligand interaction can also be determined. For example, the $K_{enhancer}$ of an enhancer of a PDZ-PL interaction can be determined, e.g., using the Kd of PDZ-PL binding as determined using the methods described *supra*. $K_{enhancer}$ is a measure of the concentration of an enhancer expected to have a biological effect. For example, administration of an enhancer of a PDZ-PL interaction in an amount sufficient to result in an intracellular inhibitor concentration of at least between about 0.1 and about 100 $K_{enhancer}$ (e.g., between about 0.5 and about 50 $K_{enhancer}$) is expected to disrupt the biological response mediated by the target PDZ-PL interaction.

Thus, in one aspect the invention provides a method of determining the potency ($K_{enhancer}$) of an enhancer or suspected enhancer of binding between a PDZ domain and a ligand by immobilizing a polypeptide comprising the PDZ domain and a non-PDZ domain on a surface, contacting the immobilized polypeptide with a plurality of different mixtures of the ligand and enhancer, wherein the different mixtures comprise a fixed amount of ligand, at least a portion of which is detectably labeled, and different concentrations of the enhancer, determining the amount of ligand bound at the different concentrations of enhancer, and calculating the potency ($K_{enhancer}$) of the enhancer from the binding based on the amount of ligand bound in the presence of different concentrations of the enhancer. Typically, at least a

portion of the ligand is detectably labeled to permit easy quantitation of ligand binding. This method, which is based on the "G" assay described *supra*, is particularly suited for high-throughput analysis of the $K_{enhancer}$ for enhancers of PDZ-ligand interactions.

It will be appreciated that the concentration of ligand and concentrations of enhancer are selected to allow meaningful detection of enhanced binding. Thus, the ligand is typically present at a concentration of between about 0.01 Kd and about 0.5 Kd and the concentrations of the test agent/enhancer typically range from 1 nM to 1 mM (e.g. a 4-fold dilution series with highest concentration 10 uM or 1 mM). In a preferred embodiment, the Kd is determined using the assay disclosed *supra*.

The potency of the binding can be determined by a variety of standard methods based on the amount of ligand bound in the presence of different concentrations of the enhancer or augmentor. For example, a plot of labeled ligand binding versus enhancer concentration can be fit to the equation:

$$S([E]) = S(0) + (S(0)*(D_{enhancer}-1)*[E]/([E]+ K_{enhancer})$$

where "$K_{enhancer}$" is the potency of the augmenting compound, and "$D_{enhancer}$" is the fold-increase in binding of the labeled ligand obtained with addition of saturating amounts of the enhancing compound, [E] is the concentration of the enhancer. It will be understood that saturating amounts are the amount of enhancer such that further addition does not significantly increase the binding signal. Knowledge of "$K_{enhancer}$" is useful because it describes a concentration of the augmenting compound in a target cell that will result in a biological effect due to dysregulation of the PDZ-PL interaction. Typical therapeutic concentrations are between about 0.1 and about 100 $K_{enhancer}$.

X. Identification of Pharmaceutical Compounds that Inhibit PDZ-PL Proteins

For certain of the PDZ proteins and PL proteins shown to bind together and for which Kd values had been obtained, additional testing was conducted to determine whether certain pharmaceutical compounds would act to antagonize or agonize the interactions. Assays were conducted as for the G' assay described supra both in the presence and absence of test compound, except that 50 ul of a 10 uM solution of the biotinylated PL peptide is allowed to react with the surface bearing the PDZ-domain polypeptide instead of a 20 uM solution as specified in step (2) of the assay.

Results from such studies are shown in **TABLES and 10A and 10B**. In

both tables, the first column (left to right) entitled "PDZ domain" lists the gene name of GST-PDZ domain fusion (see **TABLE 9**). Entries having two numbers separated by a slash indicate which PDZ domain was utilized. For example, in **TABLE 10A**, the entry for ZO-3 is 1/3. This means that PDZ domain 1 of 3 was used.

The second column labeled "PL" indicates the name of the PDZ ligand (see **TABLES 10A and 10B**) interacting with the PDZ domain. The third column entitled "Drug" lists the common or trade name of pharmaceutical compound tested and found to modulate the specific PDZ-PL interaction (suppliers and chemical information are listed in **TABLE 11**). The final column with the heading "Change in OD" indicates the change in absorbance at 450 nm of the assay in the absence (first number) or presence (second number) of chemical compound.

TABLE 11 provides the generic and commercial names for the compounds tested, as well as the Sigma Chemical Company catalog number. The molecular weight is listed in grams/mole. The final column in **TABLE 11** lists 200 times the therapeutic dose as listed in the Physicians Desk Reference and is listed in mg/ml. Stock solutions were made fresh at these concentrations and used in the assay at 10 times the therapeutic dose.

XI. Global Analysis of PDZ-PL Interactions

Certain analyses involve determining the affinity for a particular ligand and a plurality of PDZ proteins. Typically the plurality is at least 5, and often at least 25, or at least 40 different PDZ proteins. In certain analyses, the plurality of different PDZ proteins are from a particular tissue (e.g., central nervous system, spleen, cardiac muscle, kidney) or a particular class or type of cell, (e.g., a hematopoietic cell, a lymphocyte, a neuron) and the like. In some instances, the plurality of different PDZ proteins represents a substantial fraction (e.g., typically a majority, more often at least 80%) of all of the PDZ proteins known to be, or suspected of being, expressed in the tissue or cell(s), e.g., all of the PDZ proteins known to be present in lymphocytes. For example, in some analyses, the plurality is at least 50%, usually at least 80%, at least 90% or all of the PDZ proteins disclosed herein as being expressed in hematopoietic cells.

The binding of a ligand to the plurality of PDZ proteins is determined in some analyses. Using this method, it is possible to identify a particular PDZ domain bound with particular specificity by the ligand. The binding can be designated as "specific" if the affinity

of the ligand to the particular PDZ domain is at least 2-fold that of the binding to other PDZ domains in the plurality (e.g., present in that cell type). The binding is deemed "very specific" if the affinity is at least 10-fold higher than to any other PDZ in the plurality or, alternatively, at least 10-fold higher than to at least 90%, more often 95% of the other PDZs in a defined plurality. Similarly, the binding is deemed "exceedingly specific" if it is at least 100-fold higher. For example, a ligand could bind to 2 different PDZs with an affinity of 1 uM and to no other PDZs out of a set 40 with an affinity of less than 100 uM. This would constitute specific binding to those 2 PDZs. Similar measures of specificity are used to describe binding of a PDZ to a plurality of PLs.

It will be recognized that high specificity PDZ-PL interactions generally represent potentially more valuable targets for achieving a desired biological effect. The ability of an inhibitor or enhancer to act with high specificity is often desirable. In particular, the most specific PDZ-ligand interactions are also the best therapeutic targets, allowing specific inhibition of the interaction.

Identifying a high specificity interaction between a particular PDZ domain and a ligand known or suspected of binding at least one PDZ domain can be achieved with various methods. Certain methods involve providing a plurality of different immobilized polypeptides, each of said polypeptides comprising a PDZ domain and a non-PDZ domain; determining the affinity of the ligand for each of said polypeptides, and comparing the affinity of binding of the ligand to each of said polypeptides, wherein an interaction between the ligand and a particular PDZ domain is deemed to have high specificity when the ligand binds an immobilized polypeptide comprising the particular PDZ domain with at least 2-fold higher affinity than to immobilized polypeptides not comprising the particular PDZ domain.

In related methods, the affinity of binding of a specific PDZ domain to a plurality of ligands (or suspected ligands) is determined. For example, in one embodiment, the invention provides a method of identifying a high specificity interaction between a PDZ domain and a particular ligand known or suspected of binding at least one PDZ domain, by providing an immobilized polypeptide comprising the PDZ domain and a non-PDZ domain; determining the affinity of each of a plurality of ligands for the polypeptide, and comparing the affinity of binding of each of the ligands to the polypeptide, wherein an interaction between a particular ligand and the PDZ domain is deemed to have high specificity when the ligand binds an immobilized polypeptide comprising the PDZ domain with at least 2-fold higher affinity

than other ligands tested. Thus, the binding may be designated as "specific" if the affinity of the PDZ to the particular PL is at least 2-fold that of the binding to other PLs in the plurality (e.g., present in that cell type). The binding is deemed "very specific" if the affinity is at least 10-fold higher than to any other PL in the plurality or, alternatively, at least 10-fold higher than to at least 90%, more often 95% of the other PLs in a defined plurality. Similarly, the binding is deemed "exceedingly specific" if it is at least 100-fold higher. Typically the plurality is at least 5 different ligands, more often at least 10.

A. Use of Array for Global Predictions

The inventors have found that valuable information can be ascertained by analysis (e.g., simultaneous analysis) of a large number of PDZ-PL interactions. Certain analyses encompass all of the PDZ proteins expressed in a particular tissue (e.g., spleen) or type or class of cell (e.g., hematopoietic cell, neuron, lymphocyte, B cell, T cell and the like). Alternatively, the analysis encompasses at least about 5, or at least about 10, or at least about 12, or at least about 15 and often at least 50 different polypeptides, up to about 60, about 80, about 100, about 150, about 200, or even more different polypeptides; or a substantial fraction (e.g., typically a majority, more often at least 80%) of all of the PDZ proteins known to be, or suspected of being, expressed in the tissue or cell(s), e.g., all of the PDZ proteins known to be present in lymphocytes.

It will be recognized that the arrays and methods described herein are directed to the analysis of PDZ and PL interactions, and involve selection of such proteins for analysis. While the devices and methods disclosed herein can include or involve a small number of control polypeptides, they typically do not include significant numbers of proteins or fusion proteins that do not include either PDZ or PL domains (e.g., typically, at least about 90% of the arrayed or immobilized polypeptides in a method or device of the invention is a PDZ or PL sequence protein, more often at least about 95%, or at least about 99%).

It will be apparent from this disclosure that analysis of the relatively large number of different interactions preferably takes place simultaneously. In this context, "simultaneously" means that the analysis of several different PDZ-PL interactions (or the effect of a test agent on such interactions) is assessed at the same time. Typically the analysis is carried out in a high throughput (e.g., robotic) fashion. One advantage of this method of simultaneous analysis is that it permits rigorous comparison of multiple different PDZ-PL

interactions. For example, as explained in detail elsewhere herein, simultaneous analysis (and use of the arrays described *infra*) facilitates, for example, the direct comparison of the effect of an agent (e.g., an potential interaction inhibitor) on the interactions between a substantial portion of PDZs and/or PLs in a tissue or cell.

Accordingly, an array of immobilized polypeptide comprising the PDZ domain and a non-PDZ domain on a surface can be utilized in binding analyses. Typically, the array comprises at least about 5, or at least about 10, or at least about 12, or at least about 15 and often at least 50 different polypeptides. In one preferred embodiment, the different PDZ proteins are from a particular tissue (e.g., central nervous system, spleen, cardiac muscle, kidney) or a particular class or type of cell, (e.g., a hematopoietic cell, a lymphocyte, a neuron) and the like. In a most preferred embodiment, the plurality of different PDZ proteins represents a substantial fraction (e.g., typically a majority, more often at least 60%, 70% or 80%) of all of the PDZ proteins known to be, or suspected of being, expressed in the tissue or cell(s), e.g., all of the PDZ proteins known to be present in lymphocytes.

Certain arrays include a plurality, usually at least 5, 10, 25, 50 PDZ proteins present in a particular cell of interest. In this context, "array" refers to an ordered series of immobilized polypeptides in which the identity of each polypeptide is associated with its location. In some instances, the plurality of polypeptides are arrayed in a "common" area such that they can be simultaneously exposed to a solution (e.g., containing a ligand or test agent). For example, the plurality of polypeptides can be on a slide, plate or similar surface, which can be plastic, glass, metal, silica, beads or other surface to which proteins can be immobilized. In other instances, the different immobilized polypeptides are situated in separate areas, such as different wells of multi-well plate (e.g., a 24-well plate, a 96-well plate, a 384 well plate, and the like). It will be recognized that a similar advantage can be obtained by using multiple arrays in tandem.

B. Analysis of PDZ-PL Inhibition Profile

Some methods involve determining if a test compound inhibits any PDZ-ligand interaction in large set of PDZ-ligand interaction (e.g., a plurality of the PDZ-ligands interactions described in **TABLE 7** or **TABLE 12**; a majority of the PDZ-ligands identified in a particular cell or tissue as described *supra* (e.g., lymphocytes) and the like). In one embodiment, the PDZ domains of interest are expressed as GST-PDZ fusion proteins and

immobilized as described herein. For each PDZ domain, a labeled ligand that binds to the domain with a known affinity is identified as described herein.

For any known or suspected modulator (e.g., inhibitor) of a PDL-PL interaction(s), it is useful to know which interactions are inhibited (or augmented). For example, an agent that inhibits *all* PDZ-PL interactions in a cell (e.g., a lymphocyte) will have different uses than an agent that inhibits only one, or a small number, of specific PDZ-PL interactions. The profile of PDZ interactions inhibited by a particular agent is referred to as the "inhibition profile" for the agent, and is described in detail below. The profile of PDZ interactions enhanced by a particular agent is referred to as the "enhancement profile" for the agent. It will be readily apparent to one of skill guided by the description of the inhibition profile how to determine the enhancement profile for an agent. Thus, methods for determining the PDZ interaction (inhibition/enhancement) profile of an agent in a single assay are provided.

Certain methods involve determining the PDZ-PL inhibition profile of a compound by providing (i) a plurality of different immobilized polypeptides, each of said polypeptides comprising a PDZ domain and a non-PDZ domain and (ii) a plurality of corresponding ligands, wherein each ligand binds at least one PDZ domain in (i), then contacting each of said immobilized polypeptides in (i) with a corresponding ligand in (ii) in the presence and absence of a test compound, and determining for each polypeptide-ligand pair whether the test compound inhibits binding between the immobilized polypeptide and the corresponding ligand.

Typically the plurality is at least 5, and often at least 25, or at least 40 different PDZ proteins. In certain analyses, the plurality of different ligands and the plurality of different PDZ proteins are from the same tissue or a particular class or type of cell, e.g., a hematopoietic cell, a lymphocyte, a neuron and the like. In some instances, the plurality of different PDZs represents a substantial fraction (e.g., at least 80%) of all of the PDZs known to be, or suspected of being, expressed in the tissue or cell(s), e.g., all of the PDZs known to be present in lymphocytes (for example, at least 80%, at least 90% or all of the PDZs disclosed herein as being expressed in hematopoietic cells).

In certain instances, the inhibition profile is determined as follows: A plurality (e.g., all known) PDZ domains expressed in a cell (e.g., lymphocytes) are expressed as GST-fusion proteins and immobilized without altering their ligand binding properties as described *supra*. For each PDZ domain, a labeled ligand that binds to this domain with a known affinity

is identified. If the set of PDZ domains expressed in lymphocytes is denoted by {P1...Pn}, any given PDZ domain Pi binds a (labeled) ligand Li with affinity $K_d i$. To determine the inhibition profile for a test agent "compound X" the "G" assay (*supra*) can be performed as follows in 96-well plates with rows A-H and columns 1-12. Column 1 is coated with P1 and washed. The corresponding ligand L1 is added to each washed coated well of column 1 at a concentration 0.5 $K_d 1$ with (rows B, D, F, H) or without (rows A, C, E, F) between about 1 and about 1000 uM) of test compound X. Column 2 is coated with P2, and L2 (at a concentration 0.5 $K_d 2$) is added with or without inhibitor X. Additional PDZ domains and ligands are similarly tested.

Compound X is considered to inhibit the binding of Li to Pi if the average signal in the wells of column i containing X is less than half the signal in the equivalent wells of the column lacking X. Thus, in this single assay one determines the full set of lymphocyte PDZs that are inhibited by compound X.

In some embodiments, the test compound X is a mixture of compounds, such as the product of a combinatorial chemistry synthesis as described *supra*. In some embodiments, the test compound is known to have a desired biological effect, and the assay is used to determine the mechanism of action (i.e., if the biological effect is due to modulating a PDZ-PL interaction).

It will be apparent that an agent that modulates only one, or a few PDZ-PL interactions, in a panel (e.g., a panel of all known PDZs lymphocytes, a panel of at least 10, at least 20 or at least 50 PDZ domains) is a more specific modulator than an agent that modulate many or most interactions. Typically, an agent that modulates less than 20% of PDZ domains in a panel (e.g., **TABLE 7** or **TABLE 12**) is deemed a "specific" inhibitor, less than 6% a "very specific" inhibitor, and a single PDZ domain a "maximally specific" inhibitor.

It will also be appreciated that "compound X" can be a composition containing mixture of compounds (e.g., generated using combinatorial chemistry methods) rather than a single compound.

Several variations of this assay can be utilized:

In some assays, the assay above is performed using varying concentrations of the test compound X, rather than fixed concentration. This allows determination of the Ki of the X for each PDZ as described above.

In other assays, instead of pairing each PDZ Pi with a specific labeled ligand Li, a mixture of different labeled ligands is created that such that for every PDZ at least one of the

ligands in the mixture binds to this PDZ sufficiently to detect the binding in the "G" assay. This mixture is then used for every PDZ domain.

In some instances, compound X is known to have a desired biological effect, but the chemical mechanism by which it has that effect is unknown. The assays of the invention can then be used to determine if compound X has its effect by binding to a PDZ domain.

In certain assays, PDZ-domain containing proteins are classified in to groups based on their biological function, e.g. into those that regulate chemotaxis versus those that regulate transcription. An optimal inhibitor of a particular function (e.g., including but not limited to an anti-chemotactic agent, an anti-T cell activation agent, cell-cycle control, vesicle transport, apoptosis, etc.) will inhibit multiple PDZ-ligand interactions involved in the function (e.g., chemotaxis, activation) but few other interactions. Thus, the assay is used in one embodiment in screening and design of a drug that specifically blocks a particular function. For example, an agent designed to block chemotaxis might be identified because, at a given concentration, the agent inhibits 2 or more PDZs involved in chemotaxis but fewer than 3 other PDZs, or that inhibits PDZs involved in chemotaxis with a Ki > 10-fold better than for other PDZs. Thus, methods can be designed to identify an agent that inhibits a first selected PDZ-PL interaction or plurality of interactions, while not inhibiting a second selected PDZ-PL interaction or plurality of interactions. The two (or more) sets of interactions can be selected on the basis of the known biological function of the PDZ proteins, the tissue specificity of the PDZ proteins, or any other criteria. Moreover, the assay can be used to determine effective doses (i.e., drug concentrations) that result in desired biological effects while avoiding undesirable effects.

C. Side Effects of PDZ-PL Modulator Interactions

Methods can also be conducted to determine likely side effects of a therapeutic that inhibits PDZ-ligand interactions. Such methods entail identifying those target tissues, organs or cell types that express PDZ proteins and ligands that are disrupted by a specified inhibitor. If, at a therapeutic dosage, a drug intended to have an effect in one organ system (e.g., hematopoietic system) disrupts PDZ-PL interactions in a different system (e.g., CNS) it can be predicted that the drug will have effects ("side effects") on the second system. It will be apparent that the information obtained from this assay will be useful in the rational design and selection of drugs that do not have the side-effect.

In certain methods, for example, a comprehensive PDZ protein set is obtained. A "perfectly comprehensive" PDZ protein set is defined as the set of all PDZ proteins expressed in the subject animal (e.g., humans). A comprehensive set can be obtained by analysis of, for example, the human genome sequence. However, a "perfectly comprehensive" set is not required and any reasonably large set of PDZ domain proteins (e.g., the set of all known PDZ proteins; or the set listed in **TABLE 9**) will provide valuable information.

Thus, some methods involve some of all of the following steps:

a) For each PDZ protein, determine the tissues in which it is highly expressed. This can be done experimentally, although the information generally will be available in the scientific literature;

b) For each PDZ protein (or as many as possible), identify the cognate PL(s) bound by the PDZ protein;

c) Determine the Ki at which the test agent inhibits each PDZ-PL interaction, using the methods described *supra*;

d) From this information it is possible to calculate the pattern of PDZ-PL interactions disrupted at various concentrations of the test agent.

By correlating the set of PDZ-PL interactions disrupted with the expression pattern of the members of that set, it will be possible to identify the tissues likely affected by the agent.

Additional steps can also be carried out, including determining whether a specified tissue or cell type is exposed to an agent following a particular route of administration. This can be determined using basis pharmacokinetic methods and principles.

D. Modulation of Activities

The PDZ binding moieties and PDZ protein -PL protein binding antagonists of the invention are used to modulate biological activities or functions of cells (e.g., hematopoietic cells, such as T cells and B cells and the like), endothelial cells, and other immune system cells, as described herein, and for treatment of diseases and conditions in human and nonhuman animals (e.g., experimental models). Exemplary biological activities are listed *supra*.

When administered to patients, the compounds identified utilizing the methods described herein (e.g., PL-PDZ interaction inhibitors) are useful for treating (ameliorating symptoms of) a variety of diseases and conditions, including diseases characterized by inflammatory and humoral immune responses, e.g., inflammation, allergy (e.g., systemic

anaphylaxis, hypersensitivity responses, drug allergies, insect sting allergies; inflammatory bowel diseases, ulcerative colitis, ileitis and enteritis; psoriasis and inflammatory dermatoses, scleroderma; respiratory allergic diseases such as asthma, allergic rhinitis, hypersensitivity lung diseases, and the like vasculitis, rh incompatibility, transfusion reactions, drug sensitivities, PIH, atopic dermatitis, eczema, rhinnitis; autoimmune diseases, such as arthritis (rheumatoid and psoriatic), multiple sclerosis, systemic lupus erythematosus, insulin-dependent diabetes, glomerulonephritis, scleroderma, MCTD, IDDM, Hashimoto thyroiditis, Goodpasture syndrome, psoriasis and the like, osteoarthritis, polyarthritis, graft rejection (e.g., allograft rejection, e.g., renal allograft rejection, graft-vs-host disease, transplantation rejection (cardiac, kidney, lung, liver, small bowel, cornea, pancreas, cadaver, autologous, bone marrow, xenotransplantation)), atherosclerosis, angiogenesis-dependent disorders, cancers (e.g., melanomas and breast cancer, prostrate cancer, leukemias, lymphomas, metastatic disease), infectious diseases (e.g., viral infection, such as HIV, measles, parainfluenza, virus-mediated cell fusion,), ischemia (e.g., post-myocardial infarction complications, joint injury, kidney, scleroderma).

E. Agonists and Antagonists of PDZ-PL Interactions

As described herein, interactions between PDZ proteins and PL proteins in cells (e.g., hematopoietic cells, e.g., T cells and B cells) can be disrupted or inhibited by the administration of inhibitors or antagonists. Inhibitors can be identified using screening assays described herein. In some instances, the motifs disclosed herein are used to design inhibitors. In other instances, the antagonists of the invention have a structure (e.g., peptide sequence) based on the C-terminal residues of PL-domain proteins listed in **TABLE 8**. In some embodiments, the antagonists have a structure (e.g., peptide sequence) based on a PL motif disclosed herein.

The PDZ/PL antagonists and antagonists can be any of a large variety of compounds, both naturally occurring and synthetic, organic and inorganic, and including polymers (e.g., oligopeptides, polypeptides, oligonucleotides, and polynucleotides), small molecules, antibodies, sugars, fatty acids, nucleotides and nucleotide analogs, analogs of naturally occurring structures (e.g., peptide mimetics, nucleic acid analogs, and the like), and numerous other compounds. Although, for convenience, the present discussion primarily refers

antagonists of PDZ-PL interactions, it will be recognized that PDZ-PL interaction agonists can also be use in the methods disclosed herein.

In one aspect, the peptides and peptide mimetics or analogues of the invention contain an amino acid sequence that binds a PDZ domain in a cell of interest. In one embodiment, the antagonists comprise a peptide that has a sequence corresponding to the carboxy-terminal sequence of a PL protein listed in **TABLE 8**, e.g., a peptide listed **TABLE 8**. Typically, the peptide comprises at least the C-terminal two (3), three (3) or four (4) residues of the PL protein, and often the inhibitory peptide comprises more than four residues (e.g., at least five, six, seven, eight, nine, ten, twelve or fifteen residues) from the PL protein C-terminus.

In some instances, the inhibitor is a peptide, e.g., having a sequence of a PL C-terminal protein sequence.

In some embodiments, the antagonist is a fusion protein comprising such a sequence. Fusion proteins containing a transmembrane transporter amino acid sequence are particularly useful.

In other instances, the inhibitor is conserved variant of the PL C-terminal protein sequence having inhibitory activity.

In some embodiments, the antagonist is a peptide mimetic of a PL C-terminal sequence.

In some embodiments, the inhibitor is a small molecule (i.e., having a molecular weight less than 1 kD).

F. Peptide Antagonists

Certain antagonists comprise a peptide that has a sequence of a PL protein carboxy-terminus listed in **TABLE 8**. The peptide comprises at least the C-terminal two (2) residues of the PL protein, and typically, the inhibitory peptide comprises more than two residues (e.g, at least three, four, five, six, seven, eight, nine, ten, twelve or fifteen residues) from the PL protein C-terminus. The peptide can be any of a variety of lengths (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 8, at least 10, or at least 20 residues) and can contain additional residues not from the PL protein. It will be recognized that short PL peptides are sometime used in the rational design of other small molecules with similar properties.

Although most often, the residues shared by the inhibitory peptide with the PL

protein are found at the C-terminus of the peptide. However, in some embodiments, the sequence is internal. Similarly, in some cases, the inhibitory peptide comprises residues from a PL sequence that is near, but not at the c-terminus of a PL protein (see, Gee et al., 1998, *J Biological Chem.* 273:21980-87).

Sometime the PL protein carboxy-terminus sequence is referred to as the "core PDZ motif sequence" referring to the ability of the short sequence to interact with the PDZ domain. For example, in an embodiment, the "core PDZ motif sequence" contains the last four C-terminus amino acids. As described above, the four amino acid core of a PDZ motif sequence can contain additional amino acids at its amino terminus to further increase its binding affinity and/or stability. Thus, in one embodiment, the PDZ motif sequence peptide can be from four amino acids up to 15 amino acids. It is preferred that the length of the sequence to be 6-10 amino acids. More preferably, the PDZ motif sequence contains 8 amino acids. Additional amino acids at the amino terminal end of the core sequence can be derived from the natural sequence in each hematopoietic cell surface receptor or a synthetic linker. The additional amino acids can also be conservatively substituted. When the third residue from the C-terminus is S, T or Y, this residue can be phosphorylated prior to the use of the peptide.

The peptide and nonpeptide inhibitors can be small, e.g., fewer than ten amino acid residues in length if a peptide. Further, it is reported that a limited number of ligand amino acids directly contact the PDZ domain (generally less than eight) (Kozlov et al., 2000, Biochemistry 39, 2572; Doyle et al., 1996, Cell 85, 1067) and that peptides as short as the C-terminal three amino acids often retain similar binding properties to longer (> 15) amino acids peptides (Yanagisawa et al., 1997, J. Biol. Chem. 272, 8539).

G. Peptide Variants

Having identified PDZ binding peptides and PDZ-PL interaction inhibitory sequences, variations of these sequences can be made and the resulting peptide variants can be tested for PDZ domain binding or PDZ-PL inhibitory activity. In certain instances, the variants have the same or a different ability to bind a PDZ domain as the parent peptide. Typically, such amino acid substitutions are conservative, i.e., the amino acid residues are replaced with other amino acid residues having physical and/or chemical properties similar to the residues they are replacing. Preferably, conservative amino acid substitutions are those wherein an amino acid is replaced with another amino acid encompassed within the same designated class.

H. Peptide Mimetics

Having identified PDZ binding peptides and PDZ-PL interaction inhibitory sequences, peptide mimetics can be prepared using routine methods, and the inhibitory activity of the mimetics can be confirmed using the assays of the invention. Thus, certain antagonists are a peptide mimetic of a PL C-terminal sequence. The skilled artisan will recognize that individual synthetic residues and polypeptides incorporating mimetics can be synthesized using a variety of procedures and methodologies, which are well described in the scientific and patent literature, e.g., Organic Syntheses Collective Volumes, Gilman et al. (Eds) John Wiley & Sons, Inc., NY. Polypeptides incorporating mimetics can also be made using solid phase synthetic procedures, as described, e.g., by Di Marchi, et al., U.S. Pat. No. 5,422,426. Mimetics of the invention can also be synthesized using combinatorial methodologies. Various techniques for generation of peptide and peptidomimetic libraries are well known, and include, e.g., multipin, tea bag, and split-couple-mix techniques; see, e.g., al-Obeidi (1998) Mol. Biotechnol. 9:205-223; Hruby (1997) Curr. Opin. Chem. Biol. 1:114-119; Ostergaard (1997) Mol. Divers. 3:17-27; Ostresh (1996) Methods Enzymol. 267:220-234.

I. Small Molecules

In some embodiments, the inhibitor is a small molecule (i.e., having a molecular weight less than 1 kD). Methods for screening small molecules are well known in the art and include those described *supra*.

XII Preparation of Peptides

A. Chemical Synthesis

The peptides or analogues thereof that are described herein, can be prepared using virtually any art-known technique for the preparation of peptides and peptide analogues. For example, the peptides can be prepared in linear form using conventional solution or solid phase peptide syntheses and cleaved from the resin followed by purification procedures (Creighton, 1983, Protein Structures And Molecular Principles, W.H. Freeman and Co., N.Y.). Suitable procedures for synthesizing the peptides described herein are well known in the art. The composition of the synthetic peptides can be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure and mass spectroscopy).

In addition, analogues and derivatives of the peptides can be chemically synthesized. The linkage between each amino acid of the peptides of the invention can be an amide, a substituted amide or an isostere of amide. Nonclassical amino acids or chemical amino acid analogues can be introduced as a substitution or addition into the sequence. Non-classical amino acids include, but are not limited to, the D-isomers of the common amino acids, α-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ-Abu, ε-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β-methyl amino acids, Cα-methyl amino acids, Nα-methyl amino acids, and amino acid analogues in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

B. Recombinant Synthesis

If the peptide is composed entirely of gene-encoded amino acids, or a portion of it is so composed, the peptide or the relevant portion can also be synthesized using conventional recombinant genetic engineering techniques. For recombinant production, a polynucleotide sequence encoding a linear form of the peptide is inserted into an appropriate expression vehicle, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence, or in the case of an RNA viral vector, the necessary elements for replication and translation. The expression vehicle is then transfected into a suitable target cell which will express the peptide. Depending on the expression system used, the expressed peptide is then isolated by procedures well-established in the art. Methods for recombinant protein and peptide production are well known in the art (*see, e.g.*, Maniatis *et al.*, 1989, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y.; and Ausubel *et al.*, 1989, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y.).

A variety of host-expression vector systems can be utilized to express the peptides described herein. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage DNA or plasmid DNA expression vectors containing an appropriate coding sequence; yeast or filamentous fungi transformed with recombinant yeast or fungi expression vectors containing an appropriate coding sequence;

insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing an appropriate coding sequence; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus or tobacco mosaic virus) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing an appropriate coding sequence; or animal cell systems.

The expression elements of the expression systems vary in their strength and specificities. Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used in the expression vector. For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage λ, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like can be used; when cloning in insect cell systems, promoters such as the baculovirus polyhedron promoter can be used; when cloning in plant cell systems, promoters derived from the genome of plant cells (*e.g.*, heat shock promoters; the promoter for the small subunit of RUBISCO; the promoter for the chlorophyll a/b binding protein) or from plant viruses (*e.g.*, the 35S RNA promoter of CaMV; the coat protein promoter of TMV) can be used; when cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5 K promoter) can be used; when generating cell lines that contain multiple copies of expression product, SV40-, BPV- and EBV-based vectors can be used with an appropriate selectable marker.

In cases where plant expression vectors are used, the expression of sequences encoding the peptides of the invention can be driven by any of a number of promoters. For example, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV (Brisson *et al.*, 1984, Nature 310:511-514), or the coat protein promoter of TMV (Takamatsu *et al.*, 1987, EMBO J. 6:307-311) can be used; alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi *et al.*, 1984, EMBO J. 3:1671-1680; Broglie *et al.*, 1984, Science 224:838-843) or heat shock promoters, *e.g.*, soybean hsp17.5-E or hsp17.3-B (Gurley *et al.*, 1986, Mol. Cell. Biol. 6:559-565) can be used. These constructs can be introduced into planleukocytes using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, microinjection, electroporation, etc. For reviews of such techniques *see, e.g.*, Weissbach & Weissbach, 1988, Methods for Plant Molecular Biology, Academic Press, NY, Section VIII, pp. 421-463; and Grierson & Corey, 1988, Plant Molecular Biology, 2d Ed., Blackie, London, Ch. 7-9.

In one insect expression system that can be used to produce the peptides of the invention, *Autographa californica* nuclear polyhidrosis virus (AcNPV) is used as a vector to express the foreign genes. The virus grows in *Spodoptera frugiperda* cells. A coding sequence can be cloned into non-essential regions (for example the polyhedron gene) of the virus and placed under control of an AcNPV promoter (for example, the polyhedron promoter). Successful insertion of a coding sequence will result in inactivation of the polyhedron gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedron gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (*e.g.*, see Smith *et al.*, 1983, J. Virol. 46:584; Smith, U.S. Patent No. 4,215,051). Further examples of this expression system can be found in Current Protocols in Molecular Biology, Vol. 2, Ausubel *et al.*, eds., Greene Publish. Assoc. & Wiley Interscience.

In mammalian host cells, a number of viral based expression systems can be utilized. In cases where an adenovirus is used as an expression vector, a coding sequence can be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene can then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing peptide in infected hosts. (*e.g.*, See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Alternatively, the vaccinia 7.5 K promoter can be used, (*see, e.g.*, Mackett *et al.*, 1982, Proc. Natl. Acad. Sci. USA 79:7415-7419; Mackett *et al.*, 1984, J. Virol. 49:857-864; Panicali *et al.*, 1982, Proc. Natl. Acad. Sci. USA 79:4927-4931).

Other expression systems for producing linear peptides of the invention will be apparent to those having skill in the art.

Purification of the Peptides and Peptide Analogues

The peptides and peptide analogues that are provided can be purified by art-known techniques such as high performance liquid chromatography, ion exchange chromatography, gel electrophoresis, affinity chromatography and the like. The actual conditions used to purify a particular peptide or analogue will depend, in part, on factors such as net charge, hydrophobicity, hydrophilicity, etc., and will be apparent to those having skill in the art. The purified peptides can be identified by assays based on their physical or

functional properties, including radioactive labeling followed by gel electrophoresis, radioimmuno-assays, ELISA, bioassays, and the like.

For affinity chromatography purification, any antibody which specifically binds the peptides or peptide analogues can be used. For the production of antibodies, various host animals, including but not limited to rabbits, mice, rats, etc., can be immunized by injection with a peptide. The peptide can be attached to a suitable carrier, such as BSA or KLH, by means of a side chain functional group or linkers attached to a side chain functional group. Various adjuvants can be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum*.

Monoclonal antibodies to a peptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein, 1975, Nature 256:495-497, the human B-cell hybridoma technique, Kosbor *et al.*, 1983, Immunology Today 4:72; Cote *et al.*, 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030 and the EBV-hybridoma technique (Cole *et al.*, 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96 (1985)). In addition, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851-6855; Neuberger *et al.*, 1984, Nature 312:604-608; Takeda *et al.*, 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce peptide-specific single chain antibodies.

Antibody fragments which contain deletions of specific binding sites can be generated by known techniques. For example, such fragments include but are not limited to $F(ab')_2$ fragments, which can be produced by pepsin digestion of the antibody molecule and Fab fragments, which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries can be constructed (Huse *et al.*, 1989, Science

246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for the peptide of interest.

The antibody or antibody fragment specific for the desired peptide can be attached, for example, to agarose, and the antibody-agarose complex is used in immunochromatography to purify peptides of the invention. *See*, Scopes, 1984, Protein Purification: Principles and Practice, Springer-Verlag New York, Inc., NY, Livingstone, 1974, Methods Enzymology: Immunoaffinity Chromatography of Proteins 34:723-731.

XIII. Uses of PDZ Domain Binding and Antagonist Compounds

The PDZ domain-containing proteins dislcosed herein are involved in a number of biological functions, including, but not limited to, vesicular trafficking, tumor suppression, signal transduction, protein sorting, establishment of membrane polarity, apoptosis, regulation of immune response and organization of synapse formation. In general, this family of proteins has a common function of facilitating the assembly of multi-protein complexes, often serving as a bridge between several proteins, or regulating the function of other proteins. Additionally, as also noted supra, these proteins are found in essentially all cell types.

Consequently, modulation of these interactions can be utilized to control a wide variety of biological conditions and physiological conditions. In particular, modulation of interactions such as those disclosed herein can be utilized to control movement of vesicles within a cell, inhibition of tumor formation, as well as in the treatment of immune disorders, neurological disorders, muscular disorders, and intestinal disorders.

Certain compounds which modulate binding of the PDZ proteins and PL proteins can be used to inhibit leukocyte activation, which is manifested in measurable events including but not limited to, cytokine production, cell adhesion, expansion of cell numbers, apoptosis and cytotoxicity. Thus, some compounds of the invention can be used to treat diverse conditions associated with undesirable leukocyte activation, including but not limited to, acute and chronic inflammation, graft-versus-host disease, transplantation rejection, hypersensitivities and autoimmunity such as multiple sclerosis, rheumatoid arthritis, peridontal disease, systemic lupus erythematosis, juvenile diabetes mellitis, non-insulin-dependent diabetes, and allergies, and other conditions listed herein.

More specifically, in view of the various classes the PDZ and PL proteins identified herein fall into (see Section IV), the compounds can be utilized to regulate biological functions involving protein kinases, guanalyte kinases, guanine exchange factors, LIM PDZs,

tyrosine phosphatases, serine proteases, viral oncogene interacting proteins, T-cell surface receptors, B-cell surface receptors, natural killer cell receptors, monocyte surface receptors, monocyte surface receptors, granulocyte surface receptors, endothelial cell surface receptors, G-protein linked receptors, tight junction integral membrane proteins, cell adhesion molecules, neuron transport and organization molecules, regulators of G-protein signaling, ion channels and transporters and tumor associated proteins and receptors.

XIV. Formulation and Route of Administration

A. Introduction of Agonists or Antagonists (e.g., Peptides and Fusion Proteins) into Cells

In certain methods, PDZ-PL antagonists are introduced into a cell to modulate (i.e., increase or decrease) a biological function or activity of the cell. Many small organic molecules readily cross the cell membranes (or can be modified by one of skill using routine methods to increase the ability of compounds to enter cells, e.g., by reducing or eliminating charge, increasing lipophilicity, conjugating the molecule to a moiety targeting a cell surface receptor such that after interacting with the receptor). Methods for introducing larger molecules, e.g., peptides and fusion proteins are also well known, including, e.g., injection, liposome-mediated fusion, application of a hydrogel, conjugation to a targeting moiety conjugate endocytozed by the cell, electroporation, and the like).

In some instances, the antagonist or agent is a fusion polypeptide or derivatized polypeptide. A fusion or derivatized protein can include a targeting moiety that increases the ability of the polypeptide to traverse a cell membrane or causes the polypeptide to be delivered to a specified cell type (e.g., liver cells or tumor cells) preferentially or cell compartment (e.g., nuclear compartment) preferentially. Examples of targeting moieties include lipid tails, amino acid sequences such as antennapoedia peptide or a nuclear localization signal (NLS; e.g., Xenopus nucleoplasmin Robbins et al., 1991, Cell 64:615).

In certain approacheds, a peptide sequence or peptide analog, determined to inhibit a PDZ domain-PL protein binding by an assay described herein, is introduced into a cell by linking the sequence to an amino acid sequence that facilitates its transport through the plasma membrane (a "transmembrane transporter sequence"). Peptides with a desired activity can be used directly or fused to a transmembrane transporter sequence to facilitate their entry into cells. In the case of such a fusion peptide, each peptide can be fused with a heterologous

peptide at its amino terminus directly or by using a flexible polylinker such as the pentamer G-G-G-G-S repeated 1 to 3 times. Such linker has been used in constructing single chain antibodies (scFv) by being inserted between V_H and V_L (Bird et al., 1988, *Science* 242:423-426; Huston et al., 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:5979-5883). The linker is designed to enable the correct interaction between two beta-sheets forming the variable region of the single chain antibody. Other linkers which can be used include Glu-Gly-Lys-Ser-Ser-Gly-Ser-Gly-Ser-Glu-Ser-Lys-Val-Asp (Chaudhary et al., 1990, *Proc. Natl. Acad. Sci. U.S.A.* 87:1066-1070) and Lys-Glu-Ser-Gly-Ser-Val-Ser-Ser-Glu-Gln-Leu-Ala-Gln-Phe-Arg-Ser-Leu-Asp (Bird et al., 1988, *Science* 242:423-426).

A number of peptide sequences have been described in the art as capable of facilitating the entry of a peptide linked to these sequences into a cell through the plasma membrane (Derossi et al., 1998, Trends in Cell Biol. 8:84). For the purpose of this invention, such peptides are collectively referred to as transmembrane transporter peptides. Examples of these peptide include, but are not limited to, tat derived from HIV (Vives et al., 1997, *J. Biol. Chem.* 272:16010; Nagahara et al., 1998, *Nat. Med.* 4:1449), antennapedia from Drosophila (Derossi et al., 1994, *J. Biol. Chem.* 261:10444), VP22 from herpes simplex virus (Elliot and D'Hare, 1997, Cell 88:223-233), complementarity-determining regions (CDR) 2 and 3 of anti-DNA antibodies (Avrameas et al., 1998, *Proc. Natl Acad. Sci. U.S.A.*, 95:5601-5606), 70 KDa heat shock protein (Fujihara, 1999, *EMBO J.* 18:411-419) and transportan (Pooga et al., 1998, *FASEB J.* 12:67-77). In a preferred embodiment of the invention, a truncated HIV tat peptide having the sequence of GYGRKKRRQRRRG is used.

It is preferred that a transmembrane transporter sequence is fused to a hematopoietic cell surface receptor carboxyl terminal sequence at its amino-terminus with or without a linker. Generally, the C-terminus of a PDZ motif sequence (PL sequence) must be free in order to interact with a PDZ domain. The transmembrane transporter sequence can be used in whole or in part as long as it is capable of facilitating entry of the peptide into a cell.

In certain methods, a hematopoietic cell surface receptor C-terminal sequence can be used alone when it is delivered in a manner that allows its entry into cells in the absence of a transmembrane transporter sequence. For example, the peptide can be delivered in a liposome formulation or using a gene therapy approach by delivering a coding sequence for the PDZ motif alone or as a fusion molecule into a target cell.

Active compounds can also be administered via liposomes, which serve to target

the conjugates to a particular tissue, such as lymphoid tissue, or targeted selectively to infected cells, as well as increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to, e.g., a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes filled with a desired peptide or conjugate of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the selected inhibitor compositions. Liposomes for use in the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., Ann. Rev. Biophys. Bioeng. 9:467 (1980), U.S. Pat. Nos. 4,235,871, 4,501,728 and 4,837,028.

The targeting of liposomes using a variety of targeting agents is well known in the art (see, e.g., U.S. Patent Nos. 4,957,773 and 4,603,044). For targeting to the immune cells, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide or conjugate can be administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the conjugate being delivered, and the stage of the disease being treated.

In order to specifically deliver a PDZ motif sequence (PL sequence) peptide into a specific cell type, the peptide can be linked to a cell-specific targeting moiety, which include but are not limited to, ligands for diverse leukocyte surface molecules such as growth factors, hormones and cytokines, as well as antibodies or antigen-binding fragments thereof. Since a large number of cell surface receptors have been identified in leukocytes, ligands or antibodies specific for these receptors can be used as cell-specific targeting moieties. For example, interleukin-2, B7-1 (CD80), B7-2 (CD86) and CD40 or peptide fragments thereof can be used to specifically target activated T cells (The Leucocyte Antigen Facts Book, 1997, Barclay et al. (eds.), Academic Press). CD28, CTLA-4 and CD40L or peptide fragments

thereof can be used to specifically target B cells. Furthermore, Fc domains can be used to target certain Fc receptor-expressing cells such as monocytes.

Antibodies are the most versatile cell-specific targeting moieties because they can be generated against any cell surface antigen. Monoclonal antibodies have been generated against leukocyte lineage-specific markers such as certain CD antigens. Antibody variable region genes can be readily isolated from bybridoma cells by methods well known in the art. However, since antibodies are assembled between two heavy chains and two light chains, it is preferred that a scFv be used as a cell-specific targeting moiety in the present invention. Such scFv are comprised of V_H and V_L domains linked into a single polypeptide chain by a flexible linker peptide.

The PDZ motif sequence (PL sequence) can be linked to a transmembrane transporter sequence and a cell-specific targeting moiety to produce a tri-fusion molecule. This molecule can bind to a leukocyte surface molecule, passes through tbe membrane and targets PDZ domains. Alternatively, a PDZ motif sequence (PL sequence) can be linked to a cell-specific targeting moiety that binds to a surface molecule that internalizes the fusion peptide.

In another approach, microspheres of artificial polymers of mixed amino acids (proteinoids) have been used to deliver pharmaceuticals. For example, U.S. Pat. No. 4,925,673 describes drug-containing proteinoid microsphere carriers as well as methods for their preparation and use. These proteinoid microspheres are useful for the delivery of a number of active agents. Also see, U.S. Patent Nos. 5,907,030 and 6,033,884, which are incorporated herein by reference.

B. Introduction of Polynucleotides into Cells

By introducing gene sequences into cells, gene therapy can be used to treat conditions in which leukocytes are activated to result in deleterious consequences. In one embodiment, a polynucleotide that encodes a PL sequence peptide of the invention is introduced into a cell where it is expressed. The expressed peptide then inhibits the interaction of PDZ proteins and PL proteins in the cell.

Thus, in one embodiment, tbe polypeptides of the invention are expressed in a cell by introducing a nucleic acid (e.g., a DNA expression vector or mRNA) encoding the desired protein or peptide into the cell. Expression can be either constitutive or inducible

depending on the vector and choice of promoter. Methods for introduction and expression of nucleic acids into a cell are well known in the art and described herein.

In a specific embodiment, nucleic acids comprising a sequence encoding a peptide disclosed herein, are administered to a human subject. In this embodiment of the invention, the nucleic acid produces its encoded product that mediates a therapeutic effect. Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, Clinical Pharmacy 12:488-505; Wu and Wu, 1991, Biotherapy 3:87-95; Tolstoshev, 1993, Ann. Rev. Pharmacol. Toxicol. 32:573-596; Mulligan, 1993, Science 260:926-932; and Morgan and Anderson, 1993, Ann. Rev. Biochem. 62:191-217; May, 1993, TIBTECH 11(5):155-215. Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY.

In a preferred embodiment of the invention, the therapeutic composition comprises a coding sequence that is part of an expression vector. In particular, such a nucleic acid has a promoter operably linked to the coding sequence, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another specific embodiment, a nucleic acid molecule is used in which the coding sequence and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the nucleic acid (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

Delivery of the nucleic acid into a patient can be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any methods known in the art, *e.g.*, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by infection using a defective or attenuated retroviral or other viral vector (see U.S. Patent No. 4,980,286), by

direct injection of naked DNA, by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), by coating with lipids or cell-surface receptors or transfecting agents, by encapsulation in liposomes, microparticles, or microcapsules, by administering it in linkage to a peptide which is known to enter the nucleus, or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432) which can be used to target cell types specifically expressing the receptors. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180 dated April 16, 1992; WO 92/22635 dated December 23, 1992; WO92/20316 dated November 26, 1992; WO93/14188 dated July 22, 1993; WO 93/20221 dated October 14, 1993). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

In a preferred embodiment of the invention, adenoviruses as viral vectors can be used in gene therapy. Adenoviruses have the advantage of being capable of infecting non-dividing cells (Kozarsky and Wilson, 1993, Current Opinion in Genetics and Development 3:499-503). Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., 1991, Science 252:431-434; Rosenfeld et al., 1992, Cell 68:143-155; and Mastrangeli et al., 1993, J. Clin. Invest. 91:225-234. Furthermore, adenoviral vectors with modified tropism can be used for cell specific targeting (WO98/40508). Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., 1993, Proc. Soc. Exp. Biol. Med. 204:289-300).

In addition, retroviral vectors (see Miller et al., 1993, Meth. Enzymol. 217:581-599) have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The coding sequence to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., 1994, Biotherapy 6:291-302, which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., 1994, J. Clin. Invest. 93:644-651;

Kiem et al., 1994, Blood 83:1467-1473; Salmons and Gunzberg, 1993, Human Gene Therapy 4:129-141; and Grossman and Wilson, 1993, Curr. Opin. in Genetics and Devel. 3:110-114.

Another approach to gene therapy involves transferring a gene to cells in tissue culture. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration *in vivo* of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, lipofection, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see *e.g.*, Loeffler and Behr, 1993, Meth. Enzymol. 217:599-618; Cohen et al., 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92) and can be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny. In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding sequence, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

Oligonucleotides such as anti-sense RNA and DNA molecules, and ribozymes that function to inhibit the translation of a targeted mRNA, especially its C-terminus are also within the scope of the invention. Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by binding to targeted mRNA and preventing protein translation. In regard to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, *e.g.*, between -10 and +10 regions of a nucleotide sequence, are preferred.

The antisense oligonucleotide can comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine,

5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of target RNA sequences.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site can be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets can also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

The anti-sense RNA and DNA molecules and ribozymes of the invention can be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules can be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the RNA molecule. Such DNA sequences can be incorporated into a wide variety of vectors which contain suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA

constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Various modifications to the DNA molecules can be introduced as a means of increasing intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences of ribo- or deoxy- nucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

C. Other Pharmaceutical Compositions

The compounds of the invention can be administered to a subject *per se* or in the form of a sterile composition or a pharmaceutical composition. Pharmaceutical compositions comprising the compounds of the invention can be manufactured by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Pharmaceutical compositions can be formulated in conventional manner using one or more physiologically acceptable carriers, diluents, excipients or auxiliaries that facilitate processing of the active peptides or peptide analogues into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For topical administration the compounds of the invention can be formulated as solutions, gels, ointments, creams, suspensions, etc. as are well-known in the art.

Systemic formulations include those designed for administration by injection, *e.g.* subcutaneous, intravenous, intramuscular, intrathecal or intraperitoneal injection, as well as those designed for transdermal, transmucosal, oral or pulmonary administration.

For injection, the compounds of the invention can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. The solution can contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Alternatively, the compounds can be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art. This route of administration can be used to deliver the compounds to the nasal cavity.

For oral administration, the compounds can be readily formulated by combining the active peptides or peptide analogues with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. For oral solid formulations such as, for example, powders, capsules and tablets, suitable excipients include fillers such as sugars, such as lactose, sucrose, mannitol and sorbitol; cellulose preparations such as maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP); granulating agents; and binding agents. If desired, disintegrating agents can be added, such as the cross-linked polyvinylpyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

If desired, solid dosage forms can be sugar-coated or enteric-coated using standard techniques.

For oral liquid preparations such as, for example, suspensions, elixirs and solutions, suitable carriers, excipients or diluents include water, glycols, oils, alcohols, etc. Additionally, flavoring agents, preservatives, coloring agents and the like can be added.

For buccal administration, the compounds can take the form of tablets, lozenges, etc. formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit can be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator can be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds can also be formulated in rectal or vaginal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds can also be formulated as a depot preparation. Such long acting formulations can be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection.

Thus, for example, the compounds can be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Alternatively, other pharmaceutical delivery systems can be employed. Liposomes and emulsions are well known examples of delivery vehicles that can be used to deliver peptides and peptide analogues of the invention. Certain organic solvents such as dimethylsulfoxide also can be employed, although usually at the cost of greater toxicity. Additionally, the compounds can be delivered using a sustained-release system, such as semipermeable matrices of solid polymers containing the therapeutic agent. Various of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules can, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization can be employed.

As the compounds of the invention can contain charged side chains or termini, they can be included in any of the above-described formulations as the free acids or bases or as pharmaceutically acceptable salts. Pharmaceutically acceptable salts are those salts which substantially retain the biologic activity of the free bases and which are prepared by reaction with inorganic acids. Pharmaceutical salts tend to be more soluble in aqueous and other protic solvents than are the corresponding free base forms.

D. Effective Dosages

The compounds of the invention will generally be used in an amount effective to achieve the intended purpose. The compounds of the invention or pharmaceutical compositions thereof, are administered or applied in a therapeutically effective amount. By therapeutically effective amount is meant an amount effective ameliorate or prevent the symptoms, or prolong the survival of, the patient being treated. Determination of a therapeutically effective amount is well within the capabilities of those skilled in the art, especially in light of the detailed disclosure provided herein. An "inhibitory amount" or "inhibitory concentration" of a PL-PDZ binding inhibitor is an amount that reduces binding by at least about 40%, preferably at least about 50%, often at least about 70%, and even as much

as at least about 90%. Binding can be measured *in vitro* (e.g., in an A assay or G assay) or in situ.

For systemic administration, a therapeutically effective dose can be estimated initially from *in vitro* assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture. Such information can be used to more accurately determine useful doses in humans.

Initial dosages can also be estimated from *in vivo* data, *e.g.*, animal models, using techniques that are well known in the art. One having ordinary skill in the art could readily optimize administration to humans based on animal data.

Dosage amount and interval can be adjusted individually to provide plasma levels of the compounds that are sufficient to maintain therapeutic effect. Usual patient dosages for administration by injection range from about 0.1 to 5 mg/kg/day, preferably from about 0.5 to 1 mg/kg/day. Therapeutically effective serum levels can be achieved by administering multiple doses each day.

In cases of local administration or selective uptake, the effective local concentration of the compounds can not be related to plasma concentration. One having skill in the art will be able to optimize therapeutically effective local dosages without undue experimentation.

The amount of compound administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

The therapy can be repeated intermittently while symptoms detectable or even when they are not detectable. The therapy can be provided alone or in combination with other drugs. In the case of conditions associated with leukocyte activation such as transplantation rejection and autoimmunity, the drugs that can be used in combination with the compounds of the invention include, but are not limited to, steroid and non-steroid anti-inflammatory agents.

E. Toxicity

Preferably, a therapeutically effective dose of the compounds described herein will provide therapeutic benefit without causing substantial toxicity.

Toxicity of the compounds described herein can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, by determining the

LD_{50} (the dose lethal to 50% of the population) or the LD_{100} (the dose lethal to 100% of the population). The dose ratio between toxic and therapeutic effect is the therapeutic index. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a dosage range that is not toxic for use in human. The dosage of the compounds described herein lies preferably within a range of circulating concentrations that include the effective dose with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (*See, e.g.*, Fingl *et al.*, 1975, In: The Pharmacological Basis of Therapeutics, Ch.1, p.1).

EXAMPLE 1

TAT T-Cell Surface Receptor Carboxyl Terminus Fusion Peptides Inhibit T-Cell Activation

Materials And Methods

Peptide Synthesis

All peptides were chemically synthesized by standard procedures. For example, the Tat-CD3 carboxyl terminus fusion peptide, (GYGRKKRRQRRRGPPSSSSGL, SEQ ID NO:); Tat-CLASP1 carboxyl terminus fusion peptide, (GYGRKKRRQRRRGSISSSAEV, SEQ ID NO:); Tat-CLASP2 carboxyl terminus fusion peptide, (GYGRKKRRQRRRGMTSSSSVV, SEQ ID NO:); and Tat peptide, (GYGRKKRRQRRRG, SEQ ID NO:); were dissolved at 1 mM in PBS, pH 7, or dH2O. Stock MBPAc1-16 peptide, (AcASQKRPSQRHGSKYLA, SEQ ID NO:), was dissolved at 5 mM. All peptides were aliquoted and stored at –80°C until tested.

Cell Cultures

Cells were maintained and tested in RPMI 1640 media supplemented with 10% fetal calf serum (HyClone), 2 mM glutamine, 10 mM Hepes, 100 U/ml penicillin, 100 μg/ml streptomycin, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, and 50 μM beta mercaptoethanol.

T Cell Stimulation Assay

Supernatants were assayed for cytokine production following activation of T cell lines. Mouse T cell lines were stimulated using two different methods, either with antigen and antigen presenting cells or anti-mouse CD3.

Antigen-specific mouse T cells, BR4.2, were activated with the N-terminal 16 amino acid sequences of myelin basic protein (MBPAc1-16) and syngenic mouse splenocytes in 96-well plates. Mitomycin C-treated antigen presenting cells, 2×10^5 B10.BR, were added to each row of serially diluted MBPAc1-16 ranging from 0 to 200 μM. Next, 10 μM Tat-peptides or media alone was added to each row. Finally, 2×10^4 MBPAc1-16-specific T cell, pre-loaded with 10 μM Tat-peptides (see above), were added to all wells (Rabinowitz et al.,1997,. Proc. Natl. Acad. Sci. U.S.A., 94:8702-8707). Cells were activated during an overnight incubation at 5% CO2, 37°C. Cell supernatant was collected and stored at –80°C until assayed for cytokine production. The final volume was 200 μl/well.

Antibody against mouse CD3 (Pharmigen #145-2C11) was coated overnight at 4°C using 96-well flat bottom Elisa plates at a final concentration of 0.5 μg/ml, diluted in PBS. Just prior to use, plates were washed three times with 200 μl/well PBS to remove excess anti-CD3. To ensure that cells were given sufficient time to transduce Tat-peptides before activation, T cells (5×10^5 cells/ml) were pre-treated with or without 10 μM Tat-peptides for two hours at 5% CO2, 37°C and then diluted in media with or without 10 μM Tat-peptides to a final concentration of 2×10^4 cells per well in a final volume of 200 μl. Cells were then treated as described above.

Cytokine ELISA

IFNγ was measured from cell supernatants, described above, at ambient temperature using the Endogen, Inc. ELISA protocol 3. Briefly, 96-well, flat bottom, high binding ELISA plates were preincubated overnight with coating antibody (MM700). Plates were washed with 50 mM TRIS, 0.2% tween-20, pH 8 and they blocked for one hour with PBS plus 2% BSA. Washed plates were then incubated one hour with 25 μl of cell supernatant and 25 μl blocking buffer, or with 50 μl IFNγ standard. The presence of IFNγ was detected with a biotin-labeled anti-mouse IFNγ monoclonal antibody (MM700B, Endogen, Inc.,). Quantitative amounts of detection antibody are revealed with horseradish peroxidase-

conjugated streptavidin. The enzymatic, color, substrate for HRP, tetramethylbenzidine (TMB), was developed for up to 30 minutes and stopped with 1.0 M H_2SO_4. The absorbance at 450 nm was measured using a microtiter plate reader (Thermo Max, Molecular Devices) and the concentration of unknown IFNγ from cell supernatants was calculated from a standard curve generated by Softmax Pro software (Molecular Devices).

Results

Peptides containing Tat transporter sequences linked to C-terminal sequences of various PLs were testing for their ability to inhibit T cell activation. **FIGURE 1A** shows that the Tat-CD3 fusion peptide inhibits T cell activation mediated by peptide:MHC as compared to controls of Tat-peptide alone or no peptide. **FIGURE 1B** shows that Tat-CLASP2 carboxyl terminus fusion peptide inhibited T cell activation mediated by monoclonal anti-CD3 as compared to Tat-peptide alone. Tat-CLASP1 fusion peptide did not inhibit T cell activation in this experiment. These results indicate that peptides containing potential inhibitory sequences can be transported into T cells through transporter peptide such as Tat to disrupt surface receptor organization mediated by PDZ proteins. Disruption of PDZ-mediated surface receptor organization leads to blockage of T cell activation in response to antigen.

EXAMPLE 2

Generation of Eukaryotic Expression Constructs Bearing DNA Fragments that Encode PDZ Domain-Containing Genes or Portions of PDZ Domain Genes

This example describes the cloning of PDZ domain containing genes or portions of PDZ domain containing genes were into eukaryotic expression vectors in fusion with red fluorescent protein (RFP).

A. Strategy

DNA fragments corresponding to PDZ domain containing genes were generated by RT-PCR from jurkat cell line (transformed T-cells) derived RNA. Primers were designed to create restriction nuclease recognition sites at the PCR fragment's ends, to allow cloning of those fragments into the appropriate vectors. Subsequent to RT-PCR, DNA samples were submitted to agarose gel electrophoresis. Bands corresponding to the expected size were excised. DNA was extracted and treated with appropriate restriction endonuclease. DNA

samples were purified once more by gel electrophoresis, and gel extracted DNA fragments were coprecipitated and ligated with the appropriate linearized cloning vector. After transformation into *E.coli*, bacterial colonies were screened by PCR for the presence and correct orientation of insert. Positive clones were picked for large scale DNA preparation and the insert including the flanking vector sites were sequenced to ensure correct sequence of fragments and junctions between the vectors and fusion proteins.

B. Vectors:

Cloning vectors were pDsRED1-N1 (purchased from CLONTECH, # 6921-1) and pDsRED1-N1(+ATG), a derivative of pDsRED1-N1 generated by recombinant DNA technology.

DNA fragments to clone that contained the ATG-start codon were cloned into pDsRED1-N1. Fragments void of a proper translation initiation codon were cloned into pDsRED1-N-(+ATG), since this vector includes an translation initiation start codon. Vector pDsRED1-N1(+ATG) differs from pDsRED1 only with regard to the multiple cloning sites. The sequence that is unique to pDsRED1-N1(+ATG) is shown below; boundaries with pDsRED1-N1 are printed in lower case and correspond to nucleotides N 633 and N 662 in pDsRED1-N1, respectively.

5'-attGCCACCATGGGAATTCTGGATCCGGGAgat-3'

C. Deduced amino acid linker sequences:

Linker sequences between the cloned inserts and RFP vary depending on the vectors and on the restriction endonuclease used for cloning. Deduced linker amino acid sequences are listed in the table below; For some constructs, the first N-terminal and / or last C-terminal amino acid corresponds to a linker amino acid introduced by the cloning process but is not represented at that position in the corresponding gene.

Table 2

pDsRED1-N1, cloning approach: (fragment) Eco RI or Mfe I / Eco RI (vector)	PDZ domain insert C-term - LEU - GLN - SER - THR - VAL - PRO - ARG - ALA - ARG - ASP - PRO - PRO - VAL - ALA - THR - red fluorescent protein;
pDsRED1-N1(+ATG), cloning approach: (fragment) Eco RI / Eco RI (vector)	Start codon (MET) - GLY - ILE - PDZ domain gene insert - LEU - ASP - PRO - GLY - TYR - PRO - PRO - VAL - ALA - THR - red fluorescent protein;
pDsRED1-N1(+ATG), cloning approach: (fragment) Mfe I / Eco RI (vector)	Start codon (MET) - ARG - ILE - PDZ domain gene insert - LEU - ASP - PRO - GLY - TYR - PRO - PRO - VAL - ALA - THR - red fluorescent protein;

D. Constructs:

The deduced protein sequence of cloned inserts, primers used to generate DNA fragments by RT-PCR and accession # are given below for each construct. For all constructs, the fusion with RFP was carboxy terminal.

1. *Homo sapiens* Dishevelled 1 (DVL1)

Acc #: NM_004421

GI: 4758213

Cloning sites for all constructs: Eco RI / Eco RI

• Construct (N-P) [Covers the methionine start codon and extends over the C-terminal boundary of the DVL1 PDZ domain];

primers: 308 DVF and 311 DVR;

vector: pDsRED1-N1

aa 1 - aa 341

MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFKSMDQDFGV
VKEEIFDDNAKLPCFNGRVVSWLVLVEGAHSDAGSQGTDSHTDLPPPLERTGGIGDSR
SPSFQPDVASSRDGMDNETGTESMVSHRRDRARRRNREEAARTNGHPRGDRRRDVGL
PPDSASTALSSELESSSFVDSDEDDSTSRLSSSTEQSTSSRLIRKHKRRRRKQRLRQADR
ASSFSSMTDSTMSLNIITVTLNMERHHFLGICIVGQSNDRGDGGIYIGSIMKGGAVAAD
GRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSQTGPISLTVAKCWDPT

• Construct (N) [Covers the methionine start codon and extends to the N-terminal boundary of the DVL1 PDZ domain];

primers: 308 DVF and 345 DVR

vector: pDsRED1-N1

aa 1 - aa 197

MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFFKSMDQDFGV
VKEEIFDDNAKLPCFNGRVVSWLVLVEGAHSDAGSQGTDSHTDLPPPLERTGGIGDSR
SPSFQPDVASSRDGMDNETGTESMVSHRRDRARRRNREEAARTNGHPRGDRRRDVGL
PPDSASTALSSELESSSFVDSDEDG

• Construct (P) [Consists of the PDZ domain of DVL1];

primers: 344 DLF and 311 DVR;

vector: pDsRED1-N1(+ATG)

aa 246 - aa 341

SLNIITVTLNMERHHFLGICIVGQSNDRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQV
NDVNFENMSNDDAVRVLREIVSQTGPISLTVAKCWDPT

Primers:

308 DVF (N 128 - N 155) 5'-TCGGAATTCGTCGCGCCATGGCGGAGAC-3'

311 DVR (N 1004 - N 1032) 5'-GGGAATTCGGTCCCAGCACTTGGCCACAG-3'

344 DVF (N 873 - N 900) 5'-CCAGAATTCTCAACATCGTCACTGTCAC-3'

345 DVR (N713 - N744) 5'-TCGGAATTCCATCCTCGTCCGAGTCCACAAAG-3'

2. KIAA 0751 / 41.8 KD

Acc #: AB018294

GI: 3882222

Cloning sites for all constructs: (vector) Eco RI / (fragment) Mfe I

• Construct (N-J) [includes the third in frame-methionine (putative start) codon in (GI: 3882222) and extends c-terminal of the PDZ domain to the region on sequence divergency between KIAA 0751 (GI: 3882222) and hypothetical 41.8 Kd protein (AF007156 / GI: 3882222)];

primers: 318 KIF and 320 KIR;

vector: pDsRED1-N1

aa 389 - aa 803
MMYFGGHSLEEDLEWSEPQIKDSGVDTCSSTTLNEEHSHSDKHPVTWQPSKDGDRLIG
RILLNKRLKDGSVPRDSGAMLGLKVVGGKMTESGRLCAFITKVKKGSLADTVGHLRP
GDEVLEWNGRLLQGATFEEVYNIILESKPEPQVELVVSRPIGDIPRIPDSTHAQLESSSSS
FESQKMDRPSISVTSPMSPGMLRDVPQFLSGQLSIKLWFDKVGHQLIVTILGAKDLPSR
EDGRPRNPYVKIYFLPDRSDKNKRRTKTVKKTLEPKWNQTFIYSPVHRREFRERMLEIT
LWDQARVREEESEFLGEILIELETALLDDEPHWYKLQTHDVSSLPLPHPSPYMPRRQLH
GESPTRRLQRSKRISDSEVSDYDCDDGIGVVSDYRHDGRDLQSSTLSVPEQVMSSNHCS
PSGSPHRVDVIGRTT

•Construct (P) [consists of the PDZ domain of KIAA 0751 / 41.8 Kd hypothetical protein (GI: 3882222)];
primers: 341 KIF and 319 KIR.
vector pDsRED1-N1(+ATG)

aa 443 - aa 534

LKDGSVPRDSGAMLGLKVVGGKMTESGRLCAFITKVKKGSLADTVGHLRPGDEVLE
WNGRLLQGATFEEVYNIILESKPEPQVELVVSRPIA

Primers:

318 KIF (N 1366 - N 1393) 5'-AGACAATTGAGGAAATGATGTACTTTGG-3'

319 KIR (N 1830 - N 1857) 5'-GAACAATTGCAATAGGCCTTGAAACTAC-3'

320 KIR (N 2640 - N 2667) 5'-ACCCAATTGTAGTCCTTCCTATAACATC-3'

341 KIF (N 1567 - N 1593) 5'-ATAGAATTCTAAAAGATGGAAGTGTAC-3'

3. *Homo sapiens* PAR6

Acc #: AF265565

GI: 8468608

Cloning sites for all constructs: Eco RI / Eco RI

• Construct (N-P) [Covers the methionine start codon and extends over the C-terminal boundary of the PDZ domain];

primers: 322 PAF and 324 PAR;

vector: pDsRED1-N1

aa 1 - aa 251
MARPQRTPARSPDSIVEVKSKFDAEFRRFALPRASVSGFQEFSRLLRAVHQIPGLDVLL
GYTDAHGDLLPLTNDDSLHRALASGPPPLRLLVQKREADSSGLAFASNSLQRRKKGLL
LRPVAPLRTRPPLLISLPQDFRQVSSVIDVDLLPETHRRVRLHKHGSDRPLGFYIRDGMS
VRVAPQGLERVPGIFISRLVRGGLAESTGLLAVSDEILEVNGIEVAGKTLDQVTDMMV
ANSHNLIVTVKPANQR

• Construct (N) [Covers the methionine start codon and extends to the N-terminal boundary of the PDZ domain;

primers: 322 PAF and 343 PAR

vector: pDsRED1-N1

aa 1 - aa 147
MARPQRTPARSPDSIVEVKSKFDAEFRRFALPRASVSGFQEFSRLLRAVHQIPGLDVLL
GYTDAHGDLLPLTNDDSLHRALASGPPPLRLLVQKREADSSGLAFASNSLQRRKKGLL
LRPVAPLRTRPPLLISLPQDRQVSSVIDV

• Construct (P) [Consists of the PDZ domain of PAR6];

primers: 342 PAF and 324 PAR;

vector: pDsRED1-N1(+ATG)

aa 155 - aa 251
RRVRLHKHGSDRPLGFYIRDGMSVRVAPQGLERVPGIFISRLVRGGLAESTGLLAVSDE
ILEVNGIEVAGKTLDQVTDMMVANSHNLIVTVKPANQR

Primers

322 PAF (N 55 - N 82) 5'-CCCGAATTCGCCATGGCCCGGCCGCAGAG-3'

324 PAR (N 798 - N 825) 5'-CGTGAATTCGCTGGTTGGCGGGCTTGAC-3'

342 PAF (N 519 - N 548) 5'-GAGGAATTCCGACGGGTGCGGCTGCACAAG-3'

343 PAR (N 485 - N 516) 5'-GCAGAATTCCCACGTCTATGACTGAGGAAAC-3'

4. *Homo sapiens* post-synaptic density protein 95 (PSD95)

Acc #: ABU83192

GI: 3318652

Cloning sites for all constructs: Eco RI / Eco RI

Vector: pDsRED1-N1

- Construct (N-P3) [Covers the methionine start codon and extends over the C-terminal boundary of PDZ domain 3;

primers: 315 PSF and 304 PSR.

aa 1 - aa 442
MSQRPRAPRSALWLLAPPLLRWAPPLLTVLHSDLFQALLDILDYYEASLSESQKYRYQ
DEDTPPLEHSPAHLPNQANSPPVIVNTDTLEAPGYELQVNGTEGEMEYEEITLERGNSG
LGFSIAGGTDNPHIGDDPSIFITKIIPGGAAAQDGRLRVNDSILFVNEVDVREVTHSAAV
EALKEAGSIVRLYVMRRKPPAEKVMEIKLIKGPKGLGFSIAGGVGNQHIPGDNSIYVTK
IIEGGAAHKDGRLQIGDKILAVNSVGLEDVMHEDAVAALKNTYDVVYLKVAKPSNAY
LSDSYAPPDITTSYSQHLDNEISHSSYLGTDYPTAMTPTSPRRYSPVAKDLLGEEDIPRE
PRRIVIHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDQILSVNGVDLRNAS
HEQAAIALKNAGQTVTIIAQYKPEEYSR

primers:

315 PSF (N847 - N876) 5'-AGAGAATTCAGAGATATGTCCCAGAGACCAAG-3'

304 PSR (N 2161 - N 2189) 5'-CGAGAATTCTGTACTCTTCTGGTTTATAC-3'

5. *Homo sapiens* hCASK (CASK)

Acc #: AF032119

GI: 2641548

Cloning sites: Eco RI / Eco RI

- Construct (P) [Covers the PDZ domain of hCASK];

Note: The amino acid sequence homology between the human hCASK and the mouse mCASK-B is 100% identical.

primers: 336 CAF and 335 CAR;

vector: pDsRED1-N1(+ATG)

aa 399 - aa 572
RLVQFQKNTDEPMGITLKMNELNHCIVARIMHGGMIHRQGTLHVGDEIREINGISVAN
QTVEQLQKMLREMRGSITFKIVPSYR

Primers

336 CAF (N 1484 - N 1512) 5'-CCAGAATTCGGCTGGTACAGTTTCAAAAG-3'
325 CAR (N 1722 - N 1750) 5'-ACTGAATTCGGTAACTTGGCACAATCTTG-3'

6. *Homo sapiens* membrane protein, palmitolated 2 (MPP2 / DLG2)

Acc #: X82895

GI: 939884

Cloning sites for all constructs: Eco RI / Eco RI

- Construct (N-SH3) [Covers the methionine start codon, the PDZ domain and extends to the C-terminal boundary of the MPP2 SH3 domain; the construct is a splice variant of the construct annotated under GI:939884. With respect to GI:939884, the DNA portion N 238 to 309 is missing; this DNA stretch corresponds to AA 51-74. The open reading frame is maintained throughout the deletion].

primers: 305 MF and 306 MR;

vector: pDsRED1-N1

aa 1 - aa 317
MPVAATNSETAMQQVLDNLGSLPSATGAAELDLIFLRGIMESPIVRSLAKAHERLEETK
LEAVRDNNLELVQEILRDLAQLAEQSSTAAELAHILQEPHFQSLLETHDSVASKTYETP
PPSPGLDPTFSNQPVPPDAVRMVGIRKTAGEHLGVTFRVEGGELVIARILHGGMVAQQ
GLLHVGDIIKEVNGQPVGSDPRALQELLRNASGSVILKILPSYQEPHLPRQVFVKCHFD
YDPARDSLIPCKEAGLRFNAGDLLQIVNQDDANWWQACHVEGGSAGLIPSQLLEEKR
KG

Primers:

305 MF (N 58 - N 84) 5'-AGAGAATTCAGAGCCCTTGCCTCCTTC-3'

306 MR (N 798 - N 825) 5'-TGAGAATTCCTTTCCGCTTCTCCTCCAG-3'

7. *Homo sapiens* Tax interaction protein 1 (TIP-1)

Acc #: AF028823

GI: 2613001

Cloning sites: Eco RI / Bam H1

(We determined 5' start site and 5' full length sequence by 5' RACE)

- Construct (N-C);

vector: pDsRed1-N1

aa 3 - aa 125
YIPGQPVTAVVQRVEIHKLRQGENLILGFSIGGGIDQDPSQNPFSEDKTDKGIYVTRVSE
GGPAEIAGLQSGDKIMQVNGWDMTMVTHDQARKRLTKRSEEVVRLLVTRQSLQKAV
QQSML

Primer:

1318 TIP R3-1 (N 336 - N 356) 5'-CAGTCCATGCTGTCGGATCCG-3'

1317 TIP R5-1* 5'-GTCGGAATTCCCTACATCCCG-3'

*Primer 5' end corresponds to the nucleotide that is located 29 nucleotides 5' of N 1; primer sequence corresponds to sequence determined by 5' RACE; numbering corresponds to Genbank sequence entry (GI 2613001).

EXAMPLE 3

Identification of CD95 and TAX interactions with TIP-1

A. Background

Binding between these molecules was assessed using a modified ELISA. Briefly, a GST-TIP-1 fusion was produced that contained the entire PDZ domain of human TIP-1 (Insert as in EXAMPLE 2). In addition, biotinylated peptides corresponding to the C-terminal 20 amino acids of Tax and CD95 were synthesized and purified by HPLC. Binding between these entities was detected through a colorimetric assay using avidin-HRP to bind the biotin and a peroxidase substrate (G-assay, as described *supra*). By titrating the amount of peptide and protein added to these reactions, dissociation constants (Kd) were determined as an indication of relative affinity of the peptide and fusion protein association.

B. Peptide purification

Peptides representing the C-terminal 8 or 20 amino acids of CD95 and Tax were synthesized by standard FMOC chemistry and biotinylated if not used as an unlabeled competitor. Peptides were purified by reverse phase high performance liquid chromatography (HPLC) using a Vydac 218TP C18 Reversed Phase column having the dimensions of 10*25 mm, 5 um. Approximately 40 mg of peptide was dissolved in 2.0 ml of an aqueous solution of 49.9% acetonitrile and 0.1% Tri-Fluoro acetic acid (TFA). This solution was then injected into the HPLC machine through a 25 micron syringe filter (Millipore). Buffers used to get a good separation are (A) distilled water with 0.1% TFA and (B) 0.1% TFA with Acetonitrile. Gradient Segment setup is listed in **TABLE 3** below.

Table 3

Time	A	B	C	Flow rate (ml/min)
0	96%	4%	0	5.00
30	4%	96%	0	5.00

The separation occurs based on the nature of the peptides. A peptide of overall hydrophobic nature will elute off later than a peptide of a hydrophilic nature. Fractions containing the "pure" peptide were collected and checked by Mass Spectrometer (MS). Purified peptides are lyophilized for stability and stored at –80°C for later use.

C. Construction of GST-TIP-1

DNA representing the putative open reading frame of human TIP-1 was amplified by PCR and cloned into the pGEX-3X vector (Amersham-Pharmacia) to generate a GST-TIP-1 fusion vector. GST-TIP-1 protein was produced by inducing this vector with IPTG in DH5α as recommended by the Pharmacia protocol. Cells were lysed and purified by glutathione-sepharose chromatography according to manufacturer's instructions (Pharmacia). Purified protein was dialyzed against storage buffer (PBS with 25% glycerol) and stored at –20°C (short term) or –80°C (long term).

D. "G" assay for identification of interactions between peptides and fusion proteins

Reagents and materials

- Nunc Polysorp 96 well Immuno-plate (Nunc cat#62409-005)
  (Maxisorp plates have been shown to have higher background signal)
- PBS pH 7.4 (Gibco BRL cat#16777-148) or
  AVC phosphate buffered saline, 8gm NaCl, 0.29 gm KCl, 1.44 gm Na_2HPO4, 0.24gm KH_2PO4, add H2O to 1 L and pH 7.4; 0.2 micron filter
- 2% BSA/PBS (10g of bovine serum albumin, fraction V (ICN Biomedicals cat#IC15142983) into 500 ml PBS
- Goat anti-GST mAb stock @ 5 mg/ml, store at 4°C, (Amersham Pharmacia cat#27-4577-01), dilute 1:1000 in PBS, final concentration 5 ug/ml
- HRP-Streptavidin, 2.5mg/2ml stock stored at 4°C (Zymed cat#43-4323), dilute 1:2000 into 2% BSA, final concentration at 0.5 ug/ml
- Wash Buffer, 0.2% Tween 20 in 50mM Tris pH 8.0
- TMB ready to use (Dako cat#S1600)
- 1M H_2SO_4
- 12w multichannel pipettor,
- 50 ml reagent reservoirs,
- 15 ml polypropylene conical tubes

Protocol

1) Coat plate with 100 ul of 5 ug/ml goat anti GST, O/N @ 4°C
2) Dump coating antibodies out and tap dry
3) Blocking - Add 200 ul per well 2% BSA, 2 hrs at 4°C
4) Prepare proteins in 2% BSA at 5 ug/ml
   (2ml per row or per two columns)
5) 3 washes with cold PBS (must be cold through entire experiment)
   (at last wash leave PBS in wells until immediately adding next step)
6) Add proteins at 50ul per well on ice (1 to 2 hrs at 4°C)
7) Prepare Peptides in 2% BSA (2 ml/row or /columns)
8) 3 X wash with cold PBS
9) Add peptides at 50 ul per well on ice (time on / time off)
   keep on ice after last peptide has been added for 10 minutes exactly
   place at room temp for 20 minutes exactly
10) Prepare 12 ml/plate of HRP-Streptavidin (1:2000 dilution in 2%BSA)
11) 3 X wash with cold PBS
12) Add HRP-Streptavidin at 100 ul per well on ice, 20 minutes at 4°C
13) Turn on plate reader and prepare files
14) 5 X wash with Tween wash buffer, avoiding bubbles
15) Using gloves, add TMB substrate at 100 ul per well
    - incubate in dark at room temp
    - check plate periodically (5, 10, & 20 minutes)
    - take early readings, if necessary, at 650 nm (blue)
    - at 30 minutes, stop reaction with 100 ul of 1M H_2SO_4

- take final reading at 450nm (yellow)

E. Results of binding experiments

Results of peptides representing the carboxy-terminal 20 amino acids of Tax and CD95 binding to TIP-1 are shown in **FIGURE 2A**. Clearly, Tax binds GST-TIP-1 with much higher affinity than does CD95 at equivalent peptide concentrations and with equivalent amount of GST-TIP-1 fusion protein.

F. Determination of dissociation constants for proteins interacting with TIP-1

Using the protocol for the 'G' assay described above, dissociation constants were determined by titrating the amount of peptide against a set concentration of PDZ-containing protein. Kd values were determined by identifying the peptide concentration that gave half-maximal binding to the PDZ protein. Different concentrations of PDZ-containing protein were plated in order to achieve maximal peptide binding values that were less than the absorbance maximum of the ELISA plate reader. **TABLE 4** below shows the Kd values observed for the titrated reactions.

Table 4

				Tax		CD95	
PDZ	ug/ml	nm	min	OD	Kd	OD	Kd
TIP-1	0.1	450	30	3.3	0.005		
	0.3	450	30			2.6	20.0
	0.1	450	30	2.1	0.006		
	0.3	450	30			3.5	25.0
DLG1(1-2)	0.1	450	30	3.4	0.20		
	0.3	450	30			2.6	15.0
	0.1	450	30	1.6	0.13		
	0.3	450	30			2.1	20.0

Table 4 shows the Kd values in uM for the interactions between proteins and peptides in a series of 'G-Assay' experiments. Proteins on the left are GST fusions to the PDZ domain(s) of protein indicated. Numbers in parenthesis indicate the number of PDZ domains present in the fusion construct, from the amino-terminus of the first number listed to the carboxyl terminus of the second. PDZ Ligands are listed across the top of the table, representing biotinylated peptides corresponding to the carboxy-terminal 20 amino acids of each protein. The first three columns following the PDZ indicate the concentration of fusion protein plated for the G assay, followed by the wavelength and time of reading from addition of TMB substrate. 450nm indicates a reaction halted by addition of sulfuric acid and absorbance read at 450 nm. Values beneath each ligand indicate first the maximum absorbance followed by the observed Kd in uM. Numbers in the squares are the average of duplicate or

quadruplicate reactions. Blank squares indicate that the Kd for the interaction was not tested under those conditions on the same sample plate. No binding to GST alone is observed.

G. Conclusions and summary

Peptides corresponding to the PL of Tax bind TIP-1 with much higher affinity than peptides corresponding to the PL of CD95. Comparing dissociation constants (.006 uM for Tax:TIP-1, 20 uM for CD95:TIP-1), one can see that Tax can bind TIP-1 >3000-fold more strongly than CD95. This provides an explanation for potential oncogenicity of Tax. If TIP-1 is a regulator of apoptosis through binding to CD95, then upon HTLV-1 infection of lymphoid cells the Tax oncoprotein should be able to bind TIP-1 and remove it's ability to associate with CD95 at meaningful levels. If CD95 mediated apoptosis requires TIP-1, then the ability of the body to activate apoptotic pathways in HTLV-1 infected cells and hence result in a cancerous condition.

The data presented in **TABLE 4** also suggest that affinities between PDZ domains and ligands are not specific to the PDZ domain or the PL individually, but are instead specific for each unique pair. Clearly, both TIP-1 and DLG1 proteins have different dissociation constants for different ligands. Interestingly, we observe that CD95 has similar dissociation constants for both TIP-1 and DLG1. Though CD95 has similar dissociation for both pairs, Tax has different affinities for the same proteins. Hence, if a specific PL bound PDZ 'A' with 'X' Kd and PDZ 'B' with 'Y' Kd, one could not assume that another PL that bound PDZ 'A' with 'X' Kd would bind PDZ 'B' with 'Y' Kd. This shows the unique and specific nature of PDZ:PL interactions.

EXAMPLE 4

TAX and CD95 Competition for TIP-1 binding in vitro

The differing affinities of Tax and CD95 peptides for GST-TIP-1 suggest that competition between these two can be a mechanism for the oncogenicity of viral infection. Upon infection, the higher affinity of Tax could preferentially bind TIP-1 protein available in the cell, removing the TIP-1 bound to CD95 (Fas) and thereby rendering the cell less able to undergo apoptosis. In order to test this, competition experiments between Tax and CD95 for TIP-1 binding were performed using the 'G Assay', but adding additional unlabeled competitor peptide at step 9 of the 'G Assay' presented in EXAMPLE 3 section D.

FIGURES 2B and 2C show the results of these experiments. The graphs show the amount of binding of the biotinylated 20 amino acid peptide in the presence of increasing concentrations of unlabeled 8 amino acid competitor. **FIGURE 2B** shows that 20, 100, and 500 μM Tax is able to compete for binding to TIP-1 with 20 μM labeled CD95. **FIGURE 2C** shows that it takes 100-500 μM unlabeled CD95 peptide to compete for binding of 1 μM Tax to TIP-1. Taken together, a 5-fold excess of Tax is able to compete effectively for TIP-1 binding while it takes nearly a 500-fold excess of CD95 binding to interfere with the binding of Tax to TIP-1. This provides further support for the argument that Tax has a significantly higher affinity for TIP-1 than does CD95.

EXAMPLE 5

HPV E6 Oncogene and PDZ proteins

This example demonstrates the use of PL sequence motifs identified according the to the invention in the prediction of biological function in an oncogenic virus.

Human papilloma virus (HPV) infection plays a role in development of cervical carcinoma. The oncoprotein responsible for this is the early gene E6 from strains 16, 18 and 31. E6 associates with p53 and shunts this tumor suppressor into the ubiquitin proteosomal pathway to affect transformation. Using the PL motifs disclosed herein, we noted that the E6 from oncogenic strains HPV16, 18 and 31 are PDZ ligands (PLs) with the carboxy-terminal sequence of ETQ(V/L). Similarly, the E6 of oncogenic strain HPV66 has the carboxy-terminus ESTV, which also matches the consensus PDZ binding motif.

We performed an expanded search of the HPV E6 proteins and discovered HPV70 E6 fits perfectly the described PDZ consensus ETQV, identical to HPV18 and 31. We can thus predict that HPV70 is likely oncogenic on the basis that E6 is a PDZ ligand. Other HPV strains with E6 proteins that are potential PLs (based on motifs) include 63 (LYII), 66 (ESTV), 33 (ETAL), 52 (VTQV), 58 (QTQV), and 35 (ETEV). Strains 77 (QSRQ) and 80 (GSIE) can also be PLs, although the motif matches less strongly. Others, such as E6 proteins from HPV strain 57 (RTSH) and 77 (QSRQ) do not appear to be oncogenic and do not match any known consensus for PDZ binding.

To identify PDZ domains that can be bound by oncogenic HPV E6 proteins we synthesized peptides corresponding to the C-termini of several oncogenic and non-oncogenic

E6 proteins (**TABLE 8**). These were run in the 'G Assay' (EXAMPLE 3) against a variety of PDZ domains. We found that oncogenic E6 proteins with predicted PLs bound a variety of PDZ domains at varying affinities (**TABLE 7** and **TABLE 12**). In addition, non-oncogenic E6 proteins from strains 57 and 77 did not bind any of the PDZ domains tested (**TABLE** 7 and **TABLE 12** and data not shown).

Inhibitors of the interaction of the PDZ and oncogenic E6 PLs could be identified using the methods of the invention and could be useful for inhibition of E6-mediated transformation. Such inhibitors (e.g., small molecules, peptides or recombinant proteins) could be administered to patients (e.g., by local application to the vaginal vault and the uterine cervix) to treat or prevent cervical carcinoma. Diagnostic assays for oncogenic HPV are carried out using the sequences corresponding to the HPV E6 PL to design polynucleotide (e.g., PCR) or antibody probes that distinguish E6 proteins that are PLs from those that are not PLs.

EXAMPLE 6

Ability of short (>10-mer) peptides to compete with 20-mers for binding to PDZs

A. Introduction

The potential for unlabeled 8-mers and 3-mers to compete for binding with biotinylated 20-mers to PDZ domains was examined. Interactions between a PDZ domain and two or more biotinylated peptides mimicking PDZ ligands identified through the 'G Assay' were used as model interactions. Short, 3 or 8 amino acid, unlabeled peptides were synthesized by standard techniques and used at variable concentrations with a set concentration of biotinylated 20-mer. Ability of both the 3-mer and 8-mer to inhibit longer peptide binding was observed, making PDZ:PL interactions an attractive target for design of small molecule or peptide therapeutics.

B. Methods

Peptides representing the C-terminal 3, 8 or 20 amino acids of a PDZ ligand were synthesized by standard FMOC chemistry and biotinylated if not used as an unlabeled competitor. Peptides three amino acids in length were acetylated to more properly mimic the peptide bond without introducing an amino-terminal charged group. Peptides were purified by reverse phase high performance liquid chromatography (HPLC) using a Vydac 218TP C18 Reversed Phase column having the dimensions of 10*25 mm, 5 um. Approximately 40 mg of

peptide was dissolved in 2.0 ml of and aqueous solution of 49.9% acetonitrile and 0.1% Tri-Fluoro acetic acid (TFA). This solution was then injected into the HPLC machine through a 25 micron syringe filter (Millipore). Buffers used to get a good separation are (A) distilled water with 0.1% TFA and (B) 0.1% TFA with Acetonitrile. Gradient Segment setup is listed in **TABLE 5** below.

Table 5

Time	A	B	C	Flow rate (ml/min)
0	95%	5%	0	5.00
30	5%	95%	0	5.00

"Pure" fractions were collected, checked by mass spectrometry, and lyophilized (for stability). When ready to use, peptides were dissolved to 1mM concentration in PBS, pH7, or dH2O and further diluted in PBS containing 2% BSA for use in the G Assay.

PDZ domain-containing genes used in these experiments include DLG1 and PSD95:

<u>Homo sapiens Post-synaptic density-95 (PSD-95)</u>

Acc #: U83192

GI#: 3318652

Cloning sites: Bam H1 / EcoR1

- Construct (N-C);

vector: pGEX-3X

For sequence, refer to **TABLE 9** : protein spans from amino terminal end of first PDZ domain to carboxy-terminal end of third PDZ domain in frame with GST in vector.

Primer:

8PSF1 (N1150 - N1173) 5'-TCGGATCCTTGAGGGGGAGATGGA-3'

11PSR2 (N2191 – N2168) 5'-TCGGAATTCGCTATACTCTTCTGG-3'

<u>Homo sapiens Discs Large Protein, isoform 1 (DLG-1)</u>

Acc #: U13897

GI#: 475816

Cloning sites: Bam H1 / EcoR1

- Construct (N-C);

vector: pGEX-3X

For sequence, refer to **TABLE 9** : protein spans from amino terminal end of first PDZ domain to carboxy-terminal end of third PDZ domain in frame with GST in vector.

Primer:

1DF1 (N815 – N837) 5'-TCGGATCCAGGTTAATGGCTCAG -3'
3DR2 (N1850 – N1827) 5'-TCGGAATTCGACGTGACTCTTCGG -3'

DNA representing the putative open reading frames of human PSD-95 and DLG-1 were amplified by PCR and cloned into the pGEX-3X vector (Amersham-Pharmacia) to generate a GST-fusion vector. GST-fusion proteins were produced as recommended by the Pharmacia protocol by inducing this vector with IPTG in DH5α. Cells were lysed and purified by glutathione-sepharose chromatography according to manufacturer's instructions (Pharmacia). Purified protein was dialyzed against storage buffer (PBS with 25% glycerol) and stored at –20°C (short term) or –80°C (long term).

The G Assay was performed as described in EXAMPLE 3 with the exception that when a short competitor was used, 30ul of competitor peptide (at twice the final concentration) was mixed with 30ul biotinylated 20-mer (at twice the final concentration) and then added to the well.

PSD-95 and DLG-1 were incubated in the wells at 5ug/ml as described in the G Assay protocol. Biotinylated 20-mer peptides used were 20uM CLASP-2, 20uM CD46, 10uM CD95, and 10uM KV1.3 (find sequences of peptides in **TABLE 8**). Competitors (unlabeled, short peptides) tested against each of the biotinylated peptides were 50uM 8-mer of CD95, 100uM 8-mer of CD46, 50uM 8-mer of CLASP-2, and 1mM and 500uM acetylated 3-mer of CLASP-2. To deduce sequences, refer to **TABLE 8**. All absorbances were read at 450nm after stopping TMB detection reaction at 30min. Results were normalized in each group by dividing its A_{450} by the A_{450} of the PDZ / peptide binding in the absence of competitor and converting to percentage by multiplying by 100.

Results

Table 6

PDZ protein	Biotinylated 20-mer	Conc uM	Competitor	Conc uM	% binding
PSD-95	CLASP-2	20	N/A	N/A	100
			CD95 8-mer	50	92
			CD46 8-mer	100	81
			CLASP-2 8-mer	50	85
			CLASP-2 3-mer	1000	63
			CLASP-2 3-mer	500	82
	CD46	20	N/A	N/A	100
			CD95 8-mer	50	100
			CD46 8-mer	100	95
PDZ protein	Biotinylated 20-mer	Conc uM	Competitor	Conc uM	% binding
			CLASP-2 8-mer	50	90
			CLASP-2 3-mer	1000	59
			CLASP-2 3-mer	500	75
	CD95	10	N/A	N/A	100
			CD95 8-mer	50	75
			CD46 8-mer	100	65
			CLASP-2 8-mer	50	80
			CLASP-2 3-mer	1000	55
			CLASP-2 3-mer	500	55
	KV1.3	10	N/A	N/A	100
			CD95 8-mer	50	87
			CD46 8-mer	100	71
			CLASP-2 8-mer	50	82
			CLASP-2 3-mer	1000	50
			CLASP-2 3-mer	500	81
DLG-1	CLASP-2	20	N/A	N/A	100
			CD95 8-mer	50	73
			CD46 8-mer	100	90
			CLASP-2 8-mer	50	93

			CLASP-2 3-mer	1000	59
			CLASP-2 3-mer	500	61
	CD46	20	N/A	N/A	100
			CD95 8-mer	50	110
			CD46 8-mer	100	90
			CLASP-2 8-mer	50	105
			CLASP-2 3-mer	1000	45
			CLASP-2 3-mer	500	72
	CD95	10	N/A	N/A	100
			CD95 8-mer	50	70
			CD46 8-mer	100	68
			CLASP-2 8-mer	50	75
PDZ protein	Biotinylated 20-mer	Conc uM	Competitor	Conc uM	% binding
			CLASP-2 3-mer	1000	46
			CLASP-2 3-mer	500	51
	KV1.3	10	N/A	N/A	100
			CD95 8-mer	50	84
			CD46 8-mer	100	63

All standard errors are within 5% of the value.

TABLE 6 shows that it is possible to have successful competition with 3- and 8-mer unlabeled peptides against 20-mer biotinylated peptides with a 2.5-100 fold excess of unlabeled competitor. Specifically, 1mM CLASP-2 acetylated 3-mer can successfully reduce labeled ligand binding up to 50% (50-100-fold excess). With DLG-1, the 50uM CD95 8-mer can successfully reduce binding of CLASP-2 and CD95 labeled ligand approximately 30% at only 2.5 to 5-fold excess.

EXAMPLE 7

Antagonists and Agonists of PDZ/PL Interactions

A. Introduction

Many FDA approved drugs have unknown mechanism of function. It is quite possible that some of these drugs function by disrupting or increasing PDZ/PL interactions.

This possibility was examined by using the 'G Assay' (Example 3 section D). FDA approved drugs were incubated in the presence of the labeled peptide and compared to the same interaction without drug to determine if there was an effect on specific PDZ:PL interactions (drugs added with peptide at step 9 in Example 3D). The primary focus of this experiment was on drugs involved in treatment of depression (amitriptyline, desipramine, trimipramine, benztropine, and nortryptilline) and epilepsy (valproic acid). No modes of action are known for these drugs.

The FDA approved drugs used in this study are listed in **TABLE 11.** Therapeutic dose was determined by guidelines given in the Physician's Desk Reference and in the assay, 200 times this amount was used. If a dosage range was given, the upper end of the range was used. Each interaction listed in **TABLES 10A & B** was tested in the 'G Assay' (see Example 3) against each of the drugs listed in **TABLE 11.** The concentration of GST-fusion protein and peptide used in the assay represent the Kd and were determined by titration. These values can be found in **TABLE 7.** The drugs were added to the peptide before addition to the well containing the PDZ protein. Otherwise, the assay was carried out as described and read at 450nm after 30 minutes of developing. For the sequences of the PDZs and PLs used in these tests, see **TABLES 8 & 9.**

B. Results

As can be seen in **TABLE 10A**, agonist effects can be seen up to 4.3 fold higher than in the absence of drug, as in the case of AF6 and presenilin-1 in the presence of amitriptyline. Antagonistic effects have been demostrated here up to 4.2 fold higher, as in the cases of ZO-2 domain 1 and DNAM-1 in the presence of desipramine or nortryptilline and examples are listed in **TABLE 10B.**

Many agonist and antagonist effects can be seen when the drugs are incubated with PDZ/PL interactions. These results seem quite reasonable as the antidepressants used are from the tricyclic class and predominantly affect interactions where the peptide is known to function predominantly in the brain, e.g., presenilin 1 & 2 and norepinephrine transporter (NET). These results suggest that small molecules and therapeutic compounds can be used to modulate the binding between PDZ domains and their ligands.

Table 10A

Agonists 010726

PDZ domain	PL	Drug	Change in OD
ZO-3 1/3	Presenilin (115L)	Amitriptyline	1.2 to 3.6
ZO-3 1/3	Presenilin (115L)	Desipramine	1.2 to 3.3
AF6	Presenilin (115L)	Amitriptyline	0.4 to 1.7
DVL2	Presenilin (115L)	Amitriptyline	0.3 to 0.9
hSyntenin	Presenilin (115L)	Amitriptyline	1.1 to 2.7
hSyntenin	Presenilin (115L)	Desipramine	1.1 to 2.3
hSyntenin	Presenilin (115L)	Trinipramine	1.1 to 2.2
FLJ10324	Presenilin 2 (117L)	Desipramine	0.4 to 0.8
Par 3 3/3	Presenilin 2 (117L)	Desipramine	0.6 to 2.1
Mupp-1 7/13	Presenilin 2 (117L)	Desipramine	0.5 to 1.0
TIP-1 1/1	LPAP (30L)	Benztropine	1.1 to 1.6

Table 10B

Antagonists 010726

PDZ (DOMAIN)	PL	DRUG	CHANGE IN OD
ZO-1 2/3	NET (258L)	Imipramine	0.8 to 0.4
Atr-P (1/6)	DNAM (22L)	Desiprimine	4 to 1.5
BAI-1 (2/6)	DNAM (22L)	Desiprimine	4 to 1.8
ZO-2 (1/3)	DNAM (22L)	Desiprimine	2.1 to 0.5
ZO-2 (1/3)	DNAM (22L)	Nortryptilline	2.1 to 0.5
Hemba 1003117	Presenilin 2 (117L)	Valproic Acid	1.2 to 0.8
Par 3 (3/3)	Presenilin 2 (117L)	Valproic Acid	0.6 to 0.2
Mupp-1 (7/13)	Presenilin 2 (117L)	Valproic Acid	0.5 to 0.2
PTPL-1 (4/5)	Presenilin 2 (117L)	Valproic Acid	1.4 to 1.1

List of interactions and therapeutics for which a modulation of binding was observed. Concentrations at which the GST-PDZ fusion protein and labeled peptide were used can be found in Table 7 or Table 12. Concentration of drug used for each test at can be found in Table 11. 'Change in OD' shows the Absorbance of the interaction as measured by the 'G Assay' in the absence of drug at the left and the Absorbance of the interaction in the presence of drug at the right.

Table 11

Generic Name	Commercial Name	Sigma No.	Mol. Weight	Thera Dose 200x mg per mL
AMITRIPTYLINE HYDROCHLORIDE	Elavil tablets and injection	A 8404	313.9	0.66
ATROPINE SULFATE	Donnatal Elixir / Tablets	A 0257	676.8	0.0044
BENZTROPINE MESYLATE	Cogetin Injection / Tablet	B 8262	403.5	0.00428
CROMOLYN SODIUM	Gastrocrom Capsules	C 0399	512.3	0.88
DESIPRAMINE HYDROCHLORIDE	Nopramin Tablets	D 3900	302.8	1.32
Imipramine HCl		113-52-0	317	0.88
NORTRIPTYLINE HYDROCHLORIDE	PAMELOR CAPSULES	N 7261	299.8	0.11
TRIMIPRAMINE MALEATE	SURMONTIL CAPSULES	T 3146	410.5	0.44
VALPROATE SODIUM	DEPACON INJECTION	P 4543	166.2	3
VALPROIC ACID	DEPAKENE CAPSULES	P 6273	144.2	2

List of drugs used in Example 7. Therapeutic dose was determined by the Physician's Desk Reference. If a range of doses was given, the higher dose was used. In the G Assay, 200 times therapeutic dose was used, as represented in the column.

*****

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent or patent application were specifically and individually indicated to be so incorporated by reference.

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
AA01.1	Clasp-1	0	Mint 1	1,2	0	1
	Clasp-1	0	KIAA807		0	1
	Clasp-1	0	KIAA0807(S)	1	0	1
	Clasp-1	0	AIPC	1	0	1
AA02.1	Clasp-2	0	PTPL-1	2	0	1
	Clasp-2	0	PSD95	1	0	1
	Clasp-2	0	Outer Membrane	1	0	1
	Clasp-2	0	NeDLG	2	0	1
	Clasp-2	0	MUPP-1	13	0	1
	Clasp-2	0	MUPP-1	10	0	1
	Clasp-2	0	Mint 1	1,2	0	1
	Clasp-2	0	KIAA807		0	1
	Clasp-2	0	KIAA1634	2	0	1
	Clasp-2	0	KIAA1634	1	0	1
	Clasp-2	0	INADL	8	0	1
	Clasp-2	0	FLJ 10324	1	0	1
	Clasp-2	0	DLG1	2	0	1
	Clasp-2	0	DLG1	1	0	1
	Clasp-2	0	BAI-1	5	0	1
	Clasp-2	0	BAI-1	2	0	1
	Clasp-2	0	AIPC	1	0	1
AA06	CD6	0	KIAA807		0	1
	CD6	0	KIAA0807(S)	1	5	1
AA07	CD34	0	KIAA0382	1	0	1
	CD34	0	SHANK	1	0	1
	CD34	0	KIAA0147	1	0	1
	CD34	0	PTN-4	1	0	1
	CD34	0	LIM RIL	1	0	1
	CD34	0	BAI-1	6	0	1
	CD34	0	KIAA1634	5	0	1
	CD34	0	Atrophin-1 Inter. Prot.	5	0	1
AA091	GAIP (G-alpha interacting protein) RGS 19	0	KIAA1526	1	0	1
AA092	alpha-1-syntrophin	0	KIAA0807(S)	1	0	1
AA093	neurofascin (chicken)	0	ZO-2	2	0	1
	neurofascin (chicken)	0	ZO-1	2	0	1
	neurofascin (chicken)	0	ZO-1	1	0	1
	neurofascin (chicken)	0	KIAA1526	1	0	1
AA095	GluR5-2 (rat)	0	KIAA0303	1	0	1
	GluR5-2 (rat)	0	KIAA0147	1	0	2
	GluR5-2 (rat)	0	PSD95	1,2,3	0	5
	GluR5-2 (rat)	0.1	PSD95	3	1	5
	GluR5-2 (rat)	0	PSD95	1	0	3
	GluR5-2 (rat)	0	MUPP-1	10	0	2
	GluR5-2 (rat)	0	MUPP-1	11	0	1
	GluR5-2 (rat)	0.1	NeDLG	1,2	1	5
	GluR5-2 (rat)	0	NeDLG	3	0	2
	GluR5-2 (rat)	0	NeDLG	2	0	1
	GluR5-2 (rat)	0	DLG2	2	0	1
	GluR5-2 (rat)	0	DLG2	1	0	1
	GluR5-2 (rat)	0	KIAA1719	3	0	1
	GluR5-2 (rat)	0	DLG1	3	0	1
	GluR5-2 (rat)	0	DLG1	2	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	GluR5-2 (rat)	0	DLG1	1	0	2
	GluR5-2 (rat)	0	DLG1	1,2	0	5
	GluR5-2 (rat)	0.15	KIAA1634	1	0.1	5
	GluR5-2 (rat)	0.3	BAI-1	2	1	5
	GluR5-2 (rat)	0	atrophin-1 interacting Protein	1	0	1
	GluR5-2 (rat)	0	KIAA0807(S)	1	0	1
AA098L	ropporin	0	TIP1	1	0	1
AA10	CD46	0	KIAA0973	1	0	2
	CD46	0	Mint 1	2	0	1
	CD46	0	BAI-1	5	0	1
AA105	CX43 (connexin 43)	0	ZO-2	2	5	1
	CX43 (connexin 43)	0	ZO-1	2	0	2
AA106	Kir2.1 (inwardly rect. K+ channel)	0	PSD95	1,2,3	0	2
	Kir2.1 (inwardly rect. K+ channel)	0	NeDLG	1,2	0	1
	Kir2.1 (inwardly rect. K+ channel)	0	Outer Membrane	1	0	4
	Kir2.1 (inwardly rect. K+ channel)	0	DLG2	2	0	1
	Kir2.1 (inwardly rect. K+ channel)	0	DLG1	2	0	1
	Kir2.1 (inwardly rect. K+ channel)	5	DLG1	1,2	5	2
	Kir2.1 (inwardly rect. K+ channel)	0	KIAA1634	1	0	1
	Kir2.1 (inwardly rect. K+ channel)	0	atrophin-1 interacting Protein	1	0	1
AA108.1	GLUR2 (glutamate receptor 2	0	PSD95	1,2,3	0	2
	GLUR2 (glutamate receptor 2	0	NeDLG	1,2	0	2
	GLUR2 (glutamate receptor 2	0	KIAA1634	1	0	4
	GLUR2 (glutamate receptor 2	0	KIAA0807(S)	1	0	1
	GLUR2 (glutamate receptor 2	0	KIAA0147	1	0	1
	GLUR2 (glutamate receptor 2	0	ENIGMA	1	0	1
	GLUR2 (glutamate receptor 2	0	DLG2	2	0	1
	GLUR2 (glutamate receptor 2	0	DLG1	2	0	1
	GLUR2 (glutamate receptor 2	0	DLG1	1,2	0	2
	GLUR2 (glutamate receptor 2	0	AIPC	1	0	2
AA111	ephrin A2	0	KIAA0382	1	0	1
	ephrin A2	0	MUPP-1	11	0	1
	ephrin A2	0	Mint 1	2	0	1
	ephrin A2	0	KIAA1719	6	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
AA112	GluR delta-2	0	Outer Membrane	1	0	2
	GluR delta-2	0	KIAA807		0	3
	GluR delta-2	0	KIAA1526	1	5	2
	GluR delta-2	4	KIAA0807(S)	1	0.5	4
AA113	SSTR2 (somatostatin recepor 2)	0	GRIP1	7	0	1
	SSTR2 (somatostatin recepor 2)	0	KIAA0382	1	0	1
	SSTR2 (somatostatin recepor 2)	0	SHANK	1	0	1
	SSTR2 (somatostatin recepor 2)	0	Mint 1	1,2	0	1
	SSTR2 (somatostatin recepor 2)	0	Mint 1	2	0	1
	SSTR2 (somatostatin recepor 2)	0	KIAA807		0	2
	SSTR2 (somatostatin recepor 2)	0	KIAA1719	6	0	1
	SSTR2 (somatostatin recepor 2)	0	KIAA1526	1	0	1
	SSTR2 (somatostatin recepor 2)	0	KIAA0807(S)	1	0	2
AA114	GLUR7 (metabotropic glutamate receptor)	0	DLG1	2	0	1
	GLUR7 (metabotropic glutamate receptor)	0	KIAA1634	1	0	1
	GLUR7 (metabotropic glutemate receptor)	0	PAR3	3	0	2
AA115	presenilin-1	0.1	ZO-3	1	1	5
	presenilin-1	0	ZO-2	1	0	1
	presenilin-1	0	ZO-1	1	0	2
	presenilin-1	0	Unnamed Protein	2	0	1
	presenilin-1	0	TIP1	1	0	2
	presenilin-1	0	KIAA0147	3	0	1
	presenilin-1	0.2	INADL	8	5	3
	presenilin-1	0	PTPL-1	4	0	4
	presenilin-1	0	INADL	5	0	2
	presenilin-1	0.2	INADL	3	0.5	5
	presenilin-1	0.1	hSyntenin	1	5	3
	presenilin-1	0.1	HEMBA 1003117	1	0.65	5
	presenilin-1	0	MUPP-1	10	0	2
	presenilin-1	0	MUPP-1	11	0	1
	presenilin-1	0	hAPXL	1	0	1
	presenilin-1	0	P55T	1	0	1
	presenilin-1	0	NOS1	1	0	2
	presenilin-1	0.15	GRIP1	6	5	3
	presenilin-1	0.3	MUPP-1	9	0.5	5
	presenilin-1	0	MUPP-1	8	0	1
	presenilin-1	0.03	MUPP-1	7	1	5
	presenilin-1	0	MUPP-1	6	0	1
	presenilin-1	0	FLJ21687	1	0	1
	presenilin-1	0	FLJ 10324	1	0	5
	presenilin-1	0	MUPP-1	2	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	presenilin-1	0	MPP2	1	0	1
	presenilin-1	0.08	Mint 1	1,2	0.5	5
	presenilin-1	0.1	Mint 1	2	1	4
	presenilin-1	0.1	Mint 1	1	5	4
	presenilin-1	0	LIM-Mystique	1	0	2
	presenilin-1	0	LIM RIL	1	0	2
	presenilin-1	0.2	KIAA807		5	4
	presenilin-1	0.1	DVL2	1	0.5	4
	presenilin-1	0	KIAA1719	6	0	5
	presenilin-1	0	KIAA1719	5	0	3
	presenilin-1	0	CASK	1	0	2
	presenilin-1	0	KIAA1634	5	0	2
	presenilin-1	0	KIAA1634	4	0	2
	presenilin-1	0	BAI-1	2	0	2
	presenilin-1	0.2	Atrophin-1 Inter. Prot.	5	5	3
	presenilin-1	0	atrophin-1 interacting Protein	4	0	2
	presenilin-1	0	atrophin-1 interacting Protein	3	0	1
	presenilin-1	0	KIAA1222	1	5	3
	presenilin-1	0	AIPC	4	0	5
	presenilin-1	0	AIPC	1	0	5
	presenilin-1	0.1	AF6	1	0.5	3
	presenilin-1	0	PAR3	3	0	5
	presenilin-1	0	KIAA0807(S)	1	0	5
	presenilin-1	0.3	ZO-3	3	5	3
AA116	MINT-2	0	KIAA0382	1	0	1
	MINT-2	0	KIAA0300	1	0	3
	MINT-2	0	PTPL-1	4	0	4
	MINT-2	0	hSyntenin	1	0	2
	MINT-2	0	HEMBA 1003117	1	0	3
	MINT-2	0	KIAA1222	1	0	1
	MINT-2	0	MUPP-1	11	0	5
	MINT-2	0	P55T	1	0	1
	MINT-2	0	PDZK-1	4	0	1
	MINT-2	0	MUPP-1	9	0	1
	MINT-2	0	MUPP-1	7	0	1
	MINT-2	0	MUPP-1	3	0	1
	MINT-2	0	FLJ 10324	1	0	4
	MINT-2	0	MUPP-1	2	0	1
	MINT-2	0	Mint 1	1,2	0	5
	MINT-2	0	Mint 1	2	0	1
	MINT-2	0	Mint 1	1	0	2
	MINT-2	0	KIAA807		0	3
	MINT-2	0	DVL2	1	0	2
	MINT-2	0	AIPC	1	0	1
	MINT-2	0	PAR3	3	0	5
	MINT-2	0	KIAA0807(S)	1	0	4
	MINT-2	0	ZO-3	3	0	1
AA117	presenilin-2	0	ZO-1	1	0	1
	presenilin-2	0	KIAA0751(L)	1	0	1
	presenilin-2	0	KIAA0561	1	0	1
	presenilin-2	0	KIAA0300	2	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	presenilin-2	0	KIAA0300	1	0	1
	presenilin-2	8	PTPL-1	4	3	2
	presenilin-2	0	INADL	3	0	1
	presenilin-2	4	HEMBA 1003117	1	0.5	4
	presenilin-2	0	NOS1	1	0	1
	presenilin-2	0	MUPP-1	9	5	3
	presenilin-2	4	MUPP-1	7	5	2
	presenilin-2	0	MUPP-1	3	5	3
	presenilin-2	1	FLJ 10324	1	1	5
	presenilin-2	6	Mint 1	1,2	2	3
	presenilin-2	0	DVL2	1	0	1
	presenilin-2	0	Atrophin-1 Inter. Prot.	5	0	1
	presenilin-2	0	AIPC	4	0	1
	presenilin-2	0	AIPC	1	0	1
	presenilin-2	0	AF6	1	5	3
	presenilin-2	2	PAR3	3	0.5	5
	presenilin-2	0	KIAA0807(S)	1	0	1
AA118	MINT-1	0	ZO-3	1	0	2
	MINT-1	0	ZO-1	1	0	1
	MINT-1	0	KIAA0382	1	0	1
	MINT-1	0	KIAA0300	1	0	4
	MINT-1	0	INADL	8	0	1
	MINT-1	0	PTPL-1	4	0	5
	MINT-1	0.8	hSyntenin	1	5	3
	MINT-1	0	HEMBA 1003117	1	0	3
	MINT-1	0	KIAA1222	1	0	1
	MINT-1	0	MUPP-1	10	0	1
	MINT-1	0	MUPP-1	11	0	5
	MINT-1	0	NOS1	1	0	4
	MINT-1	0	PDZK-1	4	0	1
	MINT-1	0	MUPP-1	9	0	3
	MINT-1	1.5	MUPP-1	7	5	3
	MINT-1	0	MUPP-1	5	0	2
	MINT-1	0	MUPP-1	3	0	3
	MINT-1	1	FLJ 10324	1	1	5
	MINT-1	0	MUPP-1	2	0	1
	MINT-1	0	MUPP-1	1	0	1
	MINT-1	0	Mint 1	1,2	0	5
	MINT-1	2	Mint 1	1	5	3
	MINT-1	0	KIAA807		0	4
	MINT-1	0	DVL2	1	0	2
	MINT-1	0	AIPC	1	0	1
	MINT-1	0	PAR3	3	0	5
	MINT-1	0	KIAA0807(S)	1	0	5
	MINT-1	0	ZO-3	3	0	1
AA121	CD68	0	X11-beta	2	0	1
	CD68	0	SHANK	1	0	1
	CD68	0	KIAA0973	1	0	1
	CD68	0	hAPXL	1	0	2
	CD68	0	GRIP1	6	0	1
	CD68	0	FLJ 10324	1	0	1
	CD68	0	Mint 1	1,2	0	1
	CD68	0	Mint 1	2	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CD68	0	DVL2	1	0	2
	CD68	0	KIAA1719	3	0	1
	CD68	0	KIAA1719	6	0	1
	CD68	0	KIAA1634	1	0	2
	CD68	0	BAI-1	2	0	3
	CD68	0	KIAA0807(S)	1	0	5
AA123	a-actinin 2	0	rat SHANK 3	1	0	1
	a-actinin 2	1	TIP1	1	0.5	5
	a-actinin 2	0	KIAA0380	1	0	1
	a-actinin 2	0	KIAA0316	1	0	1
	a-actinin 2	2.5	TAX IP2	1	5	5
	a-actinin 2	0	Syntrophin gamma-2	1	0	1
	a-actinin 2	0	Syntrophin gamma-1	1	0	1
	a-actinin 2	0	Synt. 1 alpha	1	5	3
	a-actinin 2	0	SHANK	1	0	1
	a-actinin 2	0	KIAA0147	3	0	3
	a-actinin 2	0	KIAA0147	1	0	2
	a-actinin 2	0	PTPL-1	2	0	1
	a-actinin 2	0	INADL	3	0	1
	a-actinin 2	0	KIAA0973	1	0	2
	a-actinin 2	0	hAPXL	1	0	5
	a-actinin 2	0	Outer Membrane	1	0	1
	a-actinin 2	0	Novel PDZ	1	0	2
	a-actinin 2	0	Mint 1	1,2	0	1
	a-actinin 2	0	Mint 1	2	0	1
	a-actinin 2	0	Erbin	1	0	1
	a-actinin 2	0	ENIGMA	1	0	5
	a-actinin 2	0	LIM-Mystique	1	0	5
	a-actinin 2	0	LIM RIL	1	0	5
	a-actinin 2	2	LIM Protein	1	1	4
	a-actinin 2	0	KIAA807		0	5
	a-actinin 2	0	DLG1	2	0	1
	a-actinin 2	0	DLG1	1,2	0	1
	a-actinin 2	0	BAI-1	6	0	5
	a-actinin 2	2	KIAA1634	5	1	5
	a-actinin 2	0	BAI-1	2	0	1
	a-actinin 2	0	Atrophin-1 Inter. Prot.	5	0	5
	a-actinin 2	0	KIAA1526	1	0	1
	a-actinin 2	0	AIPC	1	0	2
	a-actinin 2	0	PAR3	3	0	1
	a-actinin 2	0	KIAA0807(S)	1	0	5
AA125	zona occludens 3 (ZO-3)	0	KIAA0382	1	0	2
	zona occludens 3 (ZO-3)	0	SHANK	1	0	1
	zona occludens 3 (ZO-3)	0	PTPL-1	2	0	2
	zona occludens 3 (ZO-3)	0	KIAA0973	1	0	2
	zona occludens 3 (ZO-3)	0	MUPP-1	13	0	2
	zona occludens 3 (ZO-3)	0	hAPXL	1	0	2
	zona occludens 3 (ZO-3)	0	Novel PDZ	1	0	1
	zona occludens 3 (ZO-3)	0	MUPP-1	9	0	1
	zona occludens 3 (ZO-3)	0	MUPP-1	7	0	1
	zona occludens 3 (ZO-3)	0	Mint 1	2	0	2
	zona occludens 3 (ZO-3)	0	LIM-Mystique	1	0	2
	zona occludens 3 (ZO-3)	0	ENIGMA	1	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	zona occludens 3 (ZO-3)	0	LIM RIL	1	0	2
	zona occludens 3 (ZO-3)	0	KIAA807		0	5
	zona occludens 3 (ZO-3)	0	KIAA1634	5	0	1
	zona occludens 3 (ZO-3)	0	BAI-1	6	0	2
	zona occludens 3 (ZO-3)	0	KIAA1526	1	0	1
	zona occludens 3 (ZO-3)	0	AIPC	1	0	1
	zona occludens 3 (ZO-3)	0	PAR3	3	0	1
	zona occludens 3 (ZO-3)	0	KIAA0807(S)	1	0	5
AA13	CD95 (fas)	0	PTPL-1	4	0	1
	CD95 (fas)	0	PTPL-1	2	5	3
	CD95 (fas)	0	Outer Membrane	1	0	1
	CD95 (fas)	0	FLJ 10324	1	0	1
	CD95 (fas)	0	DLG2	2	0	1
	CD95 (fas)	0	DLG1	2	0	1
	CD95 (fas)	0	BAI-1	5	5	3
	CD95 (fas)	0	KIAA1634	4	0	1
	CD95 (fas)	0	KIAA1634	2	0	1
	CD95 (fas)	0	KIAA1634	1	0	1
	CD95 (fas)	0	AIPC	1	0	1
AA140	KIA 1481	0	TIP1	1	0	2
	KIA 1481	0	KIAA0382	1	0	5
	KIA 1481	0	SHANK	1	0	5
	KIA 1481	0	SHANK3	1	0	3
	KIA 1481	0	EBP50	1	0	2
	KIA 1481	0	EBP50	2	0	2
	KIA 1481	0	KIAA0147	1	0	2
	KIA 1481	0	INADL	5	0	1
	KIA 1481	0	KIAA0973	1	0	2
	KIA 1481	0	KIAA1095	1	0	1
	KIA 1481	0.6	hAPXL	1	0.5	5
	KIA 1481	0	Novel PDZ	2	0	1
	KIA 1481	0	Novel PDZ	1	0	1
	KIA 1481	0	PDZK1	2,3,4	0	2
	KIA 1481	0	FLJ 00011	1	0	2
	KIA 1481	0.8	Mint 1	1,2	5	3
	KIA 1481	0	Mint 1	2	0	3
	KIA 1481	0	KIAA807		0	5
	KIA 1481	0	KIAA1634	5	0	1
	KIA 1481	0	BAI-1	8	0	2
	KIA 1481	0	BAI-1	5	5	3
	KIA 1481	0	KIAA1634	2	0	1
	KIA 1481	0	KIAA1634	1	0	2
	KIA 1481	0	BAI-1	4	0	1
	KIA 1481	0	BAI-1	2	0	2
	KIA 1481	0	KIAA1526	1	0	2
	KIA 1481	0	PDZ-73	2	0	1
	KIA 1481	0	KIAA0807(S)	1	0	5
AA147	Na+/Pi cotransporter 2	0	rat SHANK 3	1	0	4
	Na+/Pi cotransporter 2	0	ZO-2	1	0	1
	Na+/Pi cotransporter 2	0	Syntrophin gamma-2	1	0	1
	Na+/Pi cotransporter 2	0	SHANK	1	0	5
	Na+/Pi cotransporter 2	0	SHANK3	1	0	5
	Na+/Pi cotransporter 2	0	EBP50	1	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	Na+/Pi cotransporter 2	0	EBP50	2	0	2
	Na+/Pi cotransporter 2	0	INADL	8	0	1
	Na+/Pi cotransporter 2	0	PIST	1	0	1
	Na+/Pi cotransportar 2	0	KIAA0973	1	0	2
	Na+/Pi cotransporter 2	0	MUPP-1	10	0	1
	Na+/Pi cotransporter 2	0	MUPP-1	13	0	1
	Na+/Pi cotransporter 2	0	hAPXL	1	0	1
	Na+/Pi cotransportar 2	0	Outer Membrane	1	0	1
	Na+/Pi cotransporter 2	0	PDZK1	2,3,4	0	1
	Na+/Pi cotransporter 2	0	FLJ 10324	1	0	1
	Na+/Pi cotransporter 2	0	Mint 1	2	0	1
	Na+/Pi cotransporter 2	0	KIAA807		0	5
	Na+/Pi cotransporter 2	0	KIAA1526	1	0	1
	Na+/Pi cotransportar 2	0	KIAA0807(S)	1	0	5
AA148L	CFTCR (cystic fibrosis transmembrana conductance regulator)	0	SHANK	1	0	1
	CFTCR (cystic fibrosis transmambrane conductance regulator)	0	KIAA807		0	1
	CFTCR (cystic fibrosis transmembrane conductanca regulator)	0	KIAA0607(S)	1	0	2
AA152L	ActRIIA	5	PTPL-1	2	5	3
	ActRIIA	5	KIAA1634	2	5	2
AA161	MINT-3	0	KIAA0561	1	0	1
	MINT-3	0	KIAA0316	1	0	2
	MINT-3	0	KIAA0973	1	0	2
	MINT-3	0	MUPP-1	11	0	2
	MINT-3	0	MUPP-1	3	0	1
	MINT-3	0	Mint 1	1,2	0	2
	MINT-3	0	Mint 1	2	0	2
	MINT-3	0	LIM Protein	1	0	1
	MINT-3	0	KIAA807		0	1
	MINT-3	0	DVL2	1	0	1
	MINT-3	0	AF6	1	0	1
	MINT-3	0	PAR3	3	0	1
	MINT-3	0	KIAA0807(S)	1	0	1
AA169L	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	0	PTPL-1	4	0	1
	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	0	hAPXL	1	0	1
	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	0	KIAA807		0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	0	AIPC	1	0	1
	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	0	PAR3	3	0	1
	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthasa) mRNA	0	KIAA0807(S)	1	0	1
AA172	RA-GEF (ras/rap1A-assoc.-GEF)	0	KIAA0147	1	0	1
	RA-GEF (ras/rap1A-assoc.-GEF)	0	PTPL-1	2	0	4
	RA-GEF (ras/rap1A-assoc.-GEF)	0	KIAA1634	2	0	2
AA177L	c-kit receptor	0	INADL	8	0	1
	c-kit receptor	0	MUPP-1	10	0	1
	c-kit receptor	0	Mint 1	2	0	1
	c-kit receptor	0	LIM RIL	1	0	1
AA178L	PDZ-binding kinase (PBK)	0	TIP1	1	0	1
	PDZ-binding kinase (PBK)	0	Syntrophin gamma-1	1	0	1
	PDZ-binding kinasa (PBK)	0	Synt. 1 alpha	1	0	1
	PDZ-binding kinase (PBK)	6	PTPL-1	2	0.5	4
	PDZ-binding kinase (PBK)	0	PSD95	1,2,3	0	1
	PDZ-binding kinasa (PBK)	0	NeDLG	1,2	0	1
	PDZ-binding kinase (PBK)	0	DLG1	1,2	0	1
	PDZ-binding kinase (PBK)	7	KIAA1634	2	1	3
	PDZ-binding kinasa (PBK)	0	BAI-1	3	0	1
	PDZ-binding kinase (PBK)	0	Atrophin-1 Inter. Prot.	5	0	1
AA180	NMDA Glutamate Receptor 2C	0	TIP1	1	0	5
	NMDA Glutamate Receptor 2C	0	KIAA0382	1	0	1
	NMDA Glutamate Receptor 2C	0	KIAA0380	1	0	1
	NMDA Glutamate Receptor 2C	0	TAX IP2	1	0	4
	NMDA Glutamate Receptor 2C	0	Syntrophin gamma-2	1	0	2
	NMDA Glutamate Receptor 2C	0	Syntrophin gamma-1	1	0	4
	NMDA Glutamate Receptor 2C	0	Synt. 1 alpha	1	0	4
	NMDA Glutamate Receptor 2C	0	KIAA0147	3	0	1
	NMDA Glutamate Receptor 2C	0	KIAA0147	2	0	1
	NMDA Glutamate Receptor 2C	0	KIAA0147	1	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	NMDA Glutamate Receptor 2C	0	INADL	8	0	1
	NMDA Glutamate Receptor 2C	0	PTPL-1	2	0	5
	NMDA Glutamate Receptor 2C	0	PTN-4	1	0	2
	NMDA Glutamate Receptor 2C	0	INADL	5	0	1
	NMDA Glutamate Receptor 2C	0	INADL	3	0	2
	NMDA Glutamate Receptor 2C	0	PSD95	1,2,3	0	5
	NMDA Glutamate Receptor 2C	0	PSD95	3	0	2
	NMDA Glutamate Receptor 2C	0	PSD95	1	0	5
	NMDA Glutamate Receptor 2C	0	KIAA0973	1	0	1
	NMDA Glutamate Receptor 2C	0	KIAA1095	1	0	1
	NMDA Glutamate Receptor 2C	0	MUPP-1	10	0	1
	NMDA Glutamate Receptor 2C	0	MUPP-1	13	0	5
	NMDA Glutamate Receptor 2C	0	NeDLG	1,2	0	5
	NMDA Glutamate Receptor 2C	0	hAPXL	1	0	1
	NMDA Glutamate Receptor 2C	0	Outer Membrane	1	0	5
	NMDA Glutamate Receptor 2C	0	NOS1	1	0	1
	NMDA Glutamate Receptor 2C	0	NeDLG	3	0	1
	NMDA Glutamate Receptor 2C	0	NeDLG	2	0	5
	NMDA Glutamate Receptor 2C	0	NeDLG	1	0	2
	NMDA Glutamate Receptor 2C	0	MUPP-1	5	0	1
	NMDA Glutamate Receptor 2C	0	FLJ 11215	1	0	1
	NMDA Glutamate Receptor 2C	0	FLJ 00011	1	0	2
	NMDA Glutamate Receptor 2C	0	Mint 1	1,2	0	1
	NMDA Glutamate Receptor 2C	0	Mint 1	2	0	1
	NMDA Glutamate Receptor 2C	0	LIMK1	1	0	1
	NMDA Glutamate Receptor 2C	0	LIM-Mystique	1	0	4

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	NMDA Glutamate Receptor 2C	0	Erbin	1	0	4
	NMDA Glutamate Receptor 2C	0	LIM RIL	1	0	5
	NMDA Glutamate Receptor 2C	0	KIAA807		0	4
	NMDA Glutamate Receptor 2C	0	DLG2	2	0	5
	NMDA Glutamate Receptor 2C	0	DLG2	1	0	4
	NMDA Glutamate Receptor 2C	0	DLG1	2	0	5
	NMDA Glutamate Receptor 2C	0	DLG1	1	0	5
	NMDA Glutamate Receptor 2C	0	DLG1	1,2	0	5
	NMDA Glutamate Receptor 2C	0	KIAA1634	5	0	1
	NMDA Glutamate Receptor 2C	0	BAI-1	6	0	2
	NMDA Glutamate Receptor 2C	0	KIAA1634	4	0	1
	NMDA Glutamate Receptor 2C	0	BAI-1	5	0	4
	NMDA Glutamate Receptor 2C	0	KIAA1634	2	0	3
	NMDA Glutamate Receptor 2C	0	KIAA1634	1	0	5
	NMDA Glutamate Receptor 2C	0	BAI-1	4	0	3
	NMDA Glutamate Receptor 2C	0	BAI-1	3	0	1
	NMDA Glutamate Receptor 2C	0	BAI-1	2	0	4
	NMDA Glutamate Receptor 2C	0	Atrophin-1 Inter. Prot.	5	0	5
	NMDA Glutamate Receptor 2C	0	KIAA1526	1	0	3
	NMDA Glutamate Receptor 2C	0	atrophin-1 interacting Protein	3	0	2
	NMDA Glutamate Receptor 2C	0	atrophin-1 interacting Protein	1	0	5
	NMDA Glutamate Receptor 2C	0	AIPC	1	0	3
	NMDA Glutamate Receptor 2C	0	KIAA0807(S)	1	0	5
AA182L	ephrin B2	0	ZO-3	1	0	1
	ephrin B2	0	ZO-2	2	0	1
	ephrin B2	0	ZO-2	1	0	1
	ephrin B2	0	ZO-1	2	0	2
	ephrin B2	6	ZO-1	1	5	3
	ephrin B2	0	X11-beta	2	0	1
	ephrin B2	0	X11-beta	1	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	ephrin B2	0	TIP1	1	0	2
	ephrin B2	0	KIAA0382	1	0	2
	ephrin B2	0	KIAA0340	1	0	2
	ephrin B2	0	KIAA0300	1	0	2
	ephrin B2	0	Syntrophin gamma-1	1	0	2
	ephrin B2	5	SITAC-18	2	5	3
	ephrin B2	4	SITAC-18	1	5	3
	ephrin B2	0	SIP1	2	0	2
	ephrin B2	0	KIAA0147	4	0	2
	ephrin B2	0	PTPL-1	4	0	2
	ephrin B2	0	PTPL-1	2	0	2
	ephrin B2	0	INADL	3	0	2
	ephrin B2	0	PRIL16	1,2	0	2
	ephrin B2	0	hSyntenin	2	0	2
	ephrin B2	0	KIAA0973	1	0	2
	ephrin B2	0	hSyntenin	1	0	1
	ephrin B2	0	HEMBA 1003117	1	0	2
	ephrin B2	0	MUPP-1	11	0	2
	ephrin B2	0	hAPXL	1	0	1
	ephrin B2	0	Novel PDZ	1	0	2
	ephrin B2	0	NeDLG	3	0	1
	ephrin B2	0	NeDLG	2	0	2
	ephrin B2	0	PDZK-1	3	0	1
	ephrin B2	0	GRIP1	6	5	3
	ephrin B2	0	GRIP1	5	0	1
	ephrin B2	0	GRIP1	3	0	1
	ephrin B2	0	MUPP-1	6	0	2
	ephrin B2	0	MUPP-1	4	0	1
	ephrin B2	0	MUPP-1	3	0	1
	ephrin B2	0	FLJ 10324	1	0	2
	ephrin B2	0	FLJ 00011	1	0	2
	ephrin B2	0	Mint 1	1,2	0	2
	ephrin B2	0	EZRIN Phos B.P.	1	0	1
	ephrin B2	3	Mint 1	2	5	3
	ephrin B2	0	Mint 1	1	0	1
	ephrin B2	0	LIM-Mystique	1	0	1
	ephrin B2	0	LIM RIL	1	0	2
	ephrin B2	0	KIAA807		0	2
	ephrin B2	0	DLG5	2	0	1
	ephrin B2	0	DLG1	3	0	1
	ephrin B2	0	KIAA1719	5	5	4
	ephrin B2	0	CARD14	1	0	1
	ephrin B2	0	KIAA1719	1	0	1
	ephrin B2	0	BAI-1	6	0	2
	ephrin B2	0	KIAA1634	2	0	1
	ephrin B2	0	Atrophin-1 Inter. Prot.	6	0	1
	ephrin B2	0	Atrophin-1 Inter. Prot.	5	0	2
	ephrin B2	5	KIAA1526	1	5	3
	ephrin B2	0	KIAA1415	1	0	1
	ephrin B2	0	atrophin-1 interacting Protein	3	0	1
	ephrin B2	0	KIAA1284	1	0	1
	ephrin B2	0	PDZK-1	1	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	ephrin B2	0	AIPC	4	0	1
	ephrin B2	0	AIPC	3	0	1
	ephrin B2	0	AIPC	1	0	2
	ephrin B2	0	PAR3	3	0	2
	ephrin B2	0	KIAA0807(S)	1	0	2
	ephrin B2	0	ZO-3	3	0	1
	ephrin B2	0	ZO-3	2	0	2
AA183L	RhoGAP 1 (PTPL1-associated)	0	PTPL-1	4	0	2
AA185L	RGS12 (regulator of G-protein signaling 12	0	ZO-2	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	ZO-1	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	TIP1	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	PTPL-1	4	0	1
	RGS12 (regulator of G-protein signaling 12	0	PIST	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	HEMBA 1003117	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	MUPP-1	11	0	1
	RGS12 (regulator of G-protein signaling 12	0	FLJ 10324	1	0	1
	RGS12 (regulator of G-protein signaling 12	0	DLG1	1,2	0	1
	RGS12 (regulator of G-protein signaling 12	0	AF6	1	0	1
AA190L	ephrin B1	0	PTPL-1	4	0	2
	ephrin B1	0	MUPP-1	9	0	1
	ephrin B1	0	MUPP-1	7	0	1
	ephrin B1	0	MUPP-1	3	0	1
	ephrin B1	0	KIAA807		0	1
	ephrin B1	0	KIAA0807(S)	1	0	1
AA192L	JAM (junctional adhesion molecule)	0	PTPL-1	4	0	1
	JAM (junctional adhesion molecule)	0	INADL	3	0	1
	JAM (junctional adhesion molecule)	0	AF6	1	0	1
AA205L	serotonin receptor 5-HT-2C	0	INADL	8	5	1
	serotonin receptor 5-HT-2C	0	MUPP-1	10	5	1
AA206L	CITRON protein	0	TIP1	1	0	5
	CITRON protein	0	KIAA0380	1	0	1
	CITRON protein	0	Synt. 1 alpha	1	0	1
	CITRON protein	0	INADL	8	0	1
	CITRON protein	0	KIAA0973	1	0.5	5
	CITRON protein	0	MUPP-1	10	0	1
	CITRON protein	0	Outer Membrane	1	5	4
	CITRON protein	0	NeDLG	3	5	3

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CITRON protein	7	Erbin	1	5	4
	CITRON protein	0	KIAA807		0	4
	CITRON protein	0	DLG1	2	0	2
	CITRON protein	0	BAI-1	5	0	2
	CITRON protein	8	KIAA1634	4	5	3
	CITRON protein	0	KIAA1526	1	0	1
	CITRON protein	1	KIAA0807(S)	1	0.1	4
	CITRON protein	0	ZO-3	3	0	1
AA207L	Nedasin (s-form)	0	TIP1	1	0	5
	Nedasin (s-form)	0	KIAA0380	1	0	1
	Nedasin (s-form)	0	INADL	8	0	1
	Nedasin (s-form)	0	PSD95	1,2,3	0	3
	Nedasin (s-form)	0	NeDLG	1,2	0	2
	Nedasin (s-form)	0	Mint 1	1,2	0	1
	Nedasin (s-form)	0	KIAA807		0	2
	Nedasin (s-form)	0	DLG1	1,2	0	3
	Nedasin (s-form)	0	BAI-1	6	0	1
	Nedasin (s-form)	0	KIAA1634	1	0	1
	Nedasin (s-form)	0	BAI-1	2	0	1
AA210L	APC- adenomatous polyposis coli protein	0	TIP1	1	0	3
	APC- adenomatous polyposis coli protain	0	KIAA0382	1	0	1
	APC- adenometous polyposis coli protein	0	KIAA0147	1	0	1
	APC- adenomatous polyposis coli protain	0	INADL	8	0	2
	APC- adenomatous polyposis coli protein	0	PSD95	1,2,3	0	5
	APC- adenomatous polyposis coli protein	0	MUPP-1	10	0	1
	APC- adenomatous polyposis coli protein	0	NeDLG	1,2	0	4
	APC- adenomatous polyposis coli protein	0	Outar Membrana	1	0	2
	APC- edenomatous polyposis coli protein	0	FLJ 00011	1	0	1
	APC- adenomatous polyposis coli protein	0	KIAA807		0	1
	APC- adenomatous polyposis coli protein	0	DLG1	1,2	0	5
	APC- adenomatous polyposis coli protein	0	BAI-1	5	0	1
	APC- adenomatous polyposis coli protein	0	KIAA1634	2	0	1
	APC- adenomatous polyposis coli protein	0	KIAA1634	1	0	1
	APC- adenomatous polyposis coli protein	0	BAI-1	2	0	1
	APC- adenomatous polyposis coli protein	0	KIAA0807(S)	1	0	1
AA214L	ErbB-4 recaptor	0	PTPL-1	2	0	2
	ErbB-4 receptor	0	PSD95	1,2,3	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	ErbB-4 receptor	0	NeDLG	1,2	0	1
	ErbB-4 receptor	0	FLJ 10324	1	0	1
	ErbB-4 receptor	0	DLG1	1,2	0	1
	ErbB-4 receptor	0	KIAA1634	2	0	1
	ErbB-4 receptor	0	BAI-1	3	0	1
AA215	CKR5_HUMAN	0	TIP1	1	0	1
	CKR5_HUMAN	0	TAX IP2	1	0	1
	CKR5_HUMAN	0	Mint 1	1,2	0	1
	CKR5_HUMAN	0	KIAA1719	2	0	1
	CKR5_HUMAN	0	KIAA1719	5	0	1
	CKR5_HUMAN	0	KIAA1634	1	0	1
AA216	NMDA R2C	0	PTPL-1	2	0	1
	NMDA R2C	0	KIAA1634	2	0	1
AA217	catenin - delta 2	0	TIP1	1	0	3
	catenin - delta 2	0	Syntrophin gamma-1	1	0	1
	catenin - delta 2	0	KIAA0147	4	0	1
	catenin - delta 2	0	KIAA0147	2	0	3
	catenin - delta 2	0	INADL	8	0	2
	catenin - delta 2	0	PTPL-1	4	0	1
	catenin - delta 2	0	PTPL-1	2	0	5
	catenin - delta 2	0	INADL	5	0	1
	catenin - delta 2	0	PSD95	1,2,3	0	2
	catenin - delta 2	0	PSD95	1	0	1
	catenin - delta 2	0	HEMBA 1003117	1	0	1
	catenin - delta 2	0	Outer Membrane	1	0	5
	catenin - delta 2	0	NeDLG	3	0	1
	catenin - delta 2	0	FLJ 10324	1	0	3
	catenin - delta 2	0	Mint 1	1,2	0	5
	catenin - delta 2	0	Mint 1	2	0	3
	catenin - delta 2	0	Erbin	1	0	4
	catenin - delta 2	0	LIM-Mystique	1	0	5
	catenin - delta 2	0	LIM RIL	1	0	2
	catenin - delta 2	0	KIAA807		0	4
	catenin - delta 2	0	DLG2	2	0	1
	catenin - delta 2	0	DLG1	2	0	2
	catenin - delta 2	0	DLG1	1	0	1
	catenin - delta 2	0	DLG1	1,2	5	3
	catenin - delta 2	0	KIAA1634	5	0	3
	catenin - delta 2	0	BAI-1	3	0	1
	catenin - delta 2	0	Atrophin-1 Inter. Prot.	5	0	5
	catenin - delta 2	0	KIAA1526	1	0	2
	catenin - delta 2	0	atrophin-1 interacting Protein	3	0	1
	catenin - delta 2	0	AIPC	1	0	2
	catenin - delta 2	0	PAR3	3	0	1
	catenin - delta 2	0	KIAA0807(S)	1	5	3
	catenin - delta 2	0	ZO-3	3	5	3
AA218	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	GRIP1	7	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-3	1	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-2	2	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-2	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-1	2	0	4
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-1	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	X11-beta	2	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	TIP1	1	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	TIAM-2	1	0	3
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	KIAA0303	1	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	KIAA0300	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	INADL	8	0	3
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	PTPL-1	4	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	INADL	5	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	INADL	3	0	3
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	hSyntenin	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	HEMBA 1003117	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	10	0	4
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	11	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	hAPXL	1	0	3
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	Outer Membrane	1	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	NOS1	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	GRIP1	5	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	8	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	5	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	FLJ 10324	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	2	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	MUPP-1	12	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	Mint 1	1,2	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	Mint 1	2	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	Mint 1	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	LIM-Mystique	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	Erbin	1	0	3
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	LIM RIL	1	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	KIAA807		0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	DVL2	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	. 0	KIAA1719	6	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	KIAA1634	5	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	BAI-1	6	0	4
	CSPG4 (chondroitin sulfae protaoglycan 4, melanoma-associated)	0	KIAA1634	1	0	5
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	BAI-1	2	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, malanoma-associated)	0	Atrophin-1 Inter. Prot.	5	0	2
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	atrophin-1 interacting Protein	3	0	2
	CSPG4 (chondroitin sulfaa proteoglycan 4, melanoma-associated)	0	atrophin-1 interacting Protein	1	0	1
	CSPG4 (chondroitin sulfaa proteoglycan 4, malanoma-associated)	0	AIPC	1	0	5
	CSPG4 (chondroitin sulfaa protaoglycan 4, malenoma-associated)	0	AF6	1	0	5
	CSPG4 (chondroitin sulfae protaoglycan 4, melanoma-associated)	0	PAR3	3	0	3
	CSPG4 (chondroitin sulfae proteoglycen 4, melanoma-associated)	0	KIAA0807(S)	1	0	1
	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	0	ZO-3	3	0	5
AA22	DNAM-1	3	ZO-2	1	1	3
	DNAM-1	5	ZO-1	1	1	2
	DNAM-1	0	TIP1	1	0	1
	DNAM-1	5	SHANK 1	1	1	5
	DNAM-1	0	SHANK 3	1	0	2
	DNAM-1	0	EBP50	1	0	1
	DNAM-1	0	EBP50	2	0	1
	DNAM-1	0	INADL	8	0	5
	DNAM-1	2.5	PIST	1	0.5	4
	DNAM-1	2.5	MUPP-1	10	1	4
	DNAM-1	0	Outer Membrane	1	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	DNAM-1	0	NOS1	1	0	1
	DNAM-1	2	KIAA807		5	3
	DNAM-1	1	KIAA1634	1	0.3	5
	DNAM-1	4	BAI-1	2	0.1	5
	DNAM-1	3	atrophin-1 interacting Protein	1	1	3
	DNAM-1	2	KIAA0807(S)	1	5	3
AA220	claudin 10	0	DLG1	1,2	0	1
	claudin 10	0	KIAA1634	1	0	1
AA222	claudin 18	0	Mint 1	1,2	0	1
AA223	claudin 1	0	INADL	8	0	1
	claudin 1	0	Mint 1	2	0	1
AA225	claudin 9	0	Mint 1	1,2	0	1
AA226	claudin 7	0	Mint 1	1,2	5	4
AA227	claudin 2	0	Mint 1	1,2	0	2
	claudin 2	0	KIAA807		0	1
	claudin 2	0	BAI-1	3	0	1
	claudin 2	0	KIAA1634	1	0	1
AA228	Nectin 2	0	Mint 1	1,2	0	2
	Nectin 2	0	KIAA1634	1	0	1
	Nectin 2	0	AF6	1	0	2
AA23.3	Fas Ligand	0	Mint 1	1,2	0	4
	Fas Ligand	0	KIAA807		0	5
	Fas Ligand	0	KIAA0973	1	0	2
	Fas Ligand	0	KIAA0807(S)	1	0	5
	Fas Ligand	0	KIAA0380	1	0	3
	Fas Ligand	0	hAPXL	1	0	2
	Fas Ligand	0	AIPC	1	0	2
AA233L	5H2B_HUMAN	0	KIAA0316	1	0	1
	5H2B_HUMAN	0	PTPL-1	4	0	2
	5H2B_HUMAN	0.2	PTPL-1	2	0.5	5
	5H2B_HUMAN	0	PIST	1	0	1
	5H2B_HUMAN	0	HEMBA 1003117	1	0	1
	5H2B_HUMAN	0	FLJ 10324	1	0	2
	5H2B_HUMAN	0	Mint 1	1,2	5	1
	5H2B_HUMAN	0	Mint 1	2	5	1
	5H2B_HUMAN	0	KIAA807		5	1
	5H2B_HUMAN	0	KIAA1634	2	0	5
	5H2B_HUMAN	2	BAI-1	3	0.5	4
	5H2B_HUMAN	0	KIAA0807(S)	1	5	1
AA240	Dopamine transporter (Na+-dependent)	0	ZO-1	2	0	1
	Dopamine transporter (Na+-dependent)	0.4	PTPL-1	4	5	3
	Dopamine transporter (Na+-dependent)	0.3	HEMBA 1003117	1	5	5
	Dopamine transporter (Na+-dependent)	0.9	PICK1	1	5	2
	Dopamine transporter (Na+-dependent)	0.3	FLJ 10324	1	1	5
	Dopamine transporter (Na+-dependent)	0.4	KIAA807		5	4

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	Dopamine transporter (Na+-dependent)	0.9	KIAA1634	1	5	3
	Dopamine transporter (Na+-dependent)	0.4	KIAA0807(S)	1	5	4
AA243	A2AA_HUMAN (modified)	0	ZO-3	2	0	3
	A2AA_HUMAN (modified)	0	ZO-2	2	0	2
	A2AA_HUMAN (modified)	0	ZO-1	2	0	4
	A2AA_HUMAN (modified)	0	X11-beta	2	0	1
	A2AA_HUMAN (modified)	0	X11-beta	1	0	2
	A2AA_HUMAN (modified)	0	Unnamed Protein	2	0	1
	A2AA_HUMAN (modified)	0	Syntrophin gamma-1	1	0	2
	A2AA_HUMAN (modified)	0	SITAC-18	2	0	4
	A2AA_HUMAN (modified)	0	SITAC-18	1	0	4
	A2AA_HUMAN (modified)	0	PTPL-1	2	0	2
	A2AA_HUMAN (modified)	0	PAR3	3	0	2
	A2AA_HUMAN (modified)	0	MUPP-1	13	0	1
	A2AA_HUMAN (modified)	0	MUPP-1	8	0	1
	A2AA_HUMAN (modified)	0	MUPP-1	6	0	2
	A2AA_HUMAN (modified)	0	Mint 1	1	0	1
	A2AA_HUMAN (modified)	0	LIM-Mystique	1	0	1
	A2AA_HUMAN (modified)	0	KIAA1719	4	0	3
	A2AA_HUMAN (modified)	0	KIAA1526	1	0	4
	A2AA_HUMAN (modified)	0	KIAA1284	1	0	1
	A2AA_HUMAN (modified)	0	KIAA0807(S)	1	0	1
	A2AA_HUMAN (modified)	0	KIAA0751(L)	1	0	3
	A2AA_HUMAN (modified)	0	KIAA0340	1	0	1
	A2AA_HUMAN (modified)	0	INADL	4	0	1
	A2AA_HUMAN (modified)	0	INADL	3	0	2
	A2AA_HUMAN (modified)	0	HEMBA 1003117	1	0	1
	A2AA_HUMAN (modified)	0	hAPXL	1	0	1
	A2AA_HUMAN (modified)	0	FLJ21687	1	0	1
	A2AA_HUMAN (modified)	0	FLJ 10324	1	0	1
	A2AA_HUMAN (modified)	0	DLG5	2	0	1
	A2AA_HUMAN (modified)	0	CARD14	1	0	1
	A2AA_HUMAN (modified)	0	BAI-1	6	0	3
	A2AA_HUMAN (modified)	0	Atrophin-1 Inter. Prot.	6	0	1
	A2AA_HUMAN (modified)	0	Atrophin-1 Inter. Prot.	5	0	1
	A2AA_HUMAN (modified)	0	AIPC	1	0	2
AA244	A2AB_HUMAN (modified)	0	TIP1	1	0	5
	A2AB_HUMAN (modified)	0	PSD95	1,2,3	0	5
	A2AB_HUMAN (modified)	0	KIAA807		0	4
	A2AB_HUMAN (modified)	0	KIAA0303	1	0	4
	A2AB_HUMAN (modified)	0	BAI-1	4	0	5
	A2AB_HUMAN (modified)	0	BAI-1	2	0	4
AA245	A2AC_HUMAN (Modified)	0	PTPL-1	5	0	3
	A2AC_HUMAN (Modified)	0	MUPP-1	4	0	3
	A2AC_HUMAN (Modified)	0	Mint 1	2	0	3
	A2AC_HUMAN (Modified)	0	LU1	1	0	4
	A2AC_HUMAN (Modified)	0	KIAA1719	3	0	5
	A2AC_HUMAN (Modified)	0	KIAA0973	1	0	3
	A2AC_HUMAN (Modified)	0	hAPXL	1	0	3
	A2AC_HUMAN (Modified)	0	DVL2	1	0	3
	A2AC_HUMAN (Modified)	0	CARD14	1	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	A2AC_HUMAN (Modified)	0	GRIP1	5	0	1
AA248	SSR4_HUMAN	0	PDZK1	2,3,4	0	1
	SSR4_HUMAN	0	Mint 1	1,2	0	1
	SSR4_HUMAN	0	KIAA807		0	1
	SSR4_HUMAN	0	DLG1	1,2	0	1
	SSR4_HUMAN	0	BAI-1	5	0	1
	SSR4_HUMAN	0	BAI-1	4	0	1
AA25	FceRIb	0	AF6	1	0	2
	FceRIb	0	hAPXL	1	0	1
	FceRIb	0	ENIGMA	1	0	2
	FceRIb	0	LIM RIL	1	0	1
	FceRIb	0	LIM Protein	1	0	2
AA250	5-HT 3A (serotonin receptor 3A)	0	HEMBA 1003117	1	0	2
	5-HT 3A (serotonin receptor 3A)	0	MPP2	1	0	2
	5-HT 3A (serotonin receptor 3A)	0	CARD14	1	0	2
AA252	ACM3_HUMAN	0	KIAA807		0	1
	ACM3_HUMAN	0	KIAA0807(S)	1	0	1
	ACM3_HUMAN	0	hAPXL	1	0	1
	ACM3_HUMAN	0	AIPC	1	0	1
AA255	Clasp-5	0	SHANK	1	0	1
	Clasp-5	0	KIAA807		0	1
	Clasp-5	0	KIAA0807(S)	1	0	1
	Clasp-5	0	BAI-1	2	0	1
AA258	Noradrenaline transporter	0.4	ZO-1	2	5	2
	Noradrenaline transporter	1	PICK1	1	5	1
	Noradrenaline transporter	0.6	PAR3	3	1	4
	Noradrenaline transporter	0.7	MUPP-1	9	5	3
	Noradrenaline transporter	0.8	MUPP-1	7	5	3
	Noradrenaline transporter	0.4	MUPP-1	3	5	4
	Noradrenaline transporter	0.8	KIAA1719	6	5	2
	Noradrenaline transporter	0	KIAA0380	1	5	1
	Noradrenaline transporter	0.5	Mint 1	1,2	5	3
	Noradrenaline transporter	1	KIAA1719	5	5	2
	Noradrenaline transporter	0.6	INADL	3	5	3
	Noradrenaline transporter	0.6	FLJ 10324	1	5	3
	Noradrenaline transporter	0.6	AIPC	1	5	2
	Noradrenaline transporter	0.5	GRIP1	6	5	2
AA261	GABA transporter 3	0	KIAA0807(S)	1	0	1
	GABA transporter 3	0	hAPXL	1	0	1
	GABA transporter 3	0	Synt. 1 alpha	1	0	1
	GABA transporter 3	0	SHANK	1	5	1
	GABA transporter 3	0	PDZK1	2,3,4	0	1
	GABA transporter 3	0	KIAA807		0	1
AA262	Glutamate transporter 3	0	X11-beta	2	0	1
	Glutamate transporter 3	0	PTPL-1	4	5	1
	Glutamate transporter 3	0	MUPP-1	10	0	1
	Glutamate transporter 3	0	Mint 1	1,2	5	1
	Glutamate transporter 3	0	Mint 1	2	0	1
	Glutamate transporter 3	0	KIAA807		0	1
	Glutamate transporter 3	0	KIAA0807(S)	1	5	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	Glutamate transporter 3	0	hAPXL	1	0	1
	Glutamate transporter 3	0	BAI-1	4	5	1
AA264	Bone Morphogenetic Protein Receptor	0	MUPP-1	9	0	1
	Bone Morphogenetic Protein Receptor	0	MUPP-1	7	0	1
	Bone Morphogenetic Protein Receptor	0	MUPP-1	3	0	1
	Bone Morphogenetic Protein Receptor	0	KIAA0807(S)	1	0	1
AA268	PTR2_HUMAN	0	PAR3	3	0	1
	PTR2_HUMAN	0	hAPXL	1	0	1
AA269	C5AR_HUMAN	0	PTPL-1	4	0	1
AA28.1	CDW125 (modified)	0	hAPXL	1	0	1
	CDW125 (modified)	0	ENIGMA	1	0	1
AA29.2	CDw128B	0	KIAA0382	1	0	2
	CDw128B	0	SHANK	1	5	3
	CDw128B	0	KIAA807		5	5
	CDw128B	0	KIAA0807(S)	1	0	5
AA29.3	IL-8RB	0	TIP1	1	0	1
	IL-8RB	0	Synt. 1 alpha	1	0	1
	IL-8RB	0	PDZK1	2,3,4	0	1
	IL-8RB	0	Novel PDZ	2	0	1
	IL-8RB	0	MUPP-1	13	0	1
	IL-8RB	0	KIAA1634	5	0	1
	IL-8RB	0	KIAA1634	1	0	1
	IL-8RB	0	KIAA0380	1	0	1
	IL-8RB	0	BAI-1	6	0	1
	IL-8RB	0	BAI-1	2	0	1
AA30	LPAP	0	Unnamed Protein	2	0	3
	LPAP	0	KIAA0382	1	0	5
	LPAP	0	KIAA0316	1	0	1
	LPAP	0	SHANK	1	0	3
	LPAP	0	SHANK3	1	0	3
	LPAP	0	EBP50	1	0	5
	LPAP	0	EBP50	2	0	4
	LPAP	0	KIAA0147	1	0	3
	LPAP	0	PTPL-1	2	0	1
	LPAP	0	PIST	1	0	1
	LPAP	0	HEMBA 1003117	1	0	1
	LPAP	0	hAPXL	1	0	1
	LPAP	0	NOS1	1	0	1
	LPAP	0	PDZK1	2,3,4	0	3
	LPAP	0	GRIP1	3	0	1
	LPAP	0	FLJ 10324	1	0	1
	LPAP	1.5	FLJ 00011	1	5	4
	LPAP	0	Mint 1	2	0	1
	LPAP	0	KIAA807		0	5
	LPAP	0	BAI-1	2	0	2
	LPAP	0	Atrophin-1 Inter. Prot.	5	0	2
	LPAP	0	KIAA1526	1	0	1
AA300	Traf2	0	KIAA807		0	2
	Traf2	0	KIAA0973	1	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	Traf2	0	KIAA0807(S)	1	0	4
AA31	Mannose receptor	0	hAPXL	1	0	1
	Mannose receptor	0	FLJ 00011	1	0	1
	Mannose receptor	0	KIAA807		0	1
	Mannose receptor	0	KIAA0807(S)	1	5	1
AA36	Neuroligin	0	ZO-1	1	0	1
	Neuroligin	0	TIP1	1	0	1
	Neuroligin	0.3	SHANK	1	5	2
	Neuroligin	0	SHANK3	1	0	3
	Neuroligin	0	EBP50	1	0	2
	Neuroligin	0	EBP50	2	0	1
	Neuroligin	0	INADL	8	0	1
	Neuroligin	0	PTPL-1	4	0	1
	Neuroligin	0	PTPL-1	2	0	1
	Neuroligin	0	PSD95	1,2,3	0	2
	Neuroligin	0	NeDLG	1,2	0	1
	Neuroligin	0	NOS1	1	0	1
	Neuroligin	0	NeDLG	3	0	1
	Neuroligin	0	FLJ 10324	1	0	1
	Neuroligin	0	Mint 1	1,2	0	1
	Neuroligin	0	KIAA807		0	3
	Neuroligin	0	DLG1	1,2	0	2
	Neuroligin	0	KIAA1634	2	0	2
	Neuroligin	0.1	KIAA1634	1	1	4
	Neuroligin	0.25	atrophin-1 interacting Protein	1	5	2
AA37	Glycophorin C	0	KIAA1719	6	5	1
	Glycophorin C	0	PAR3	3	0	2
AA40	Dock2	0	KIAA0382	1	0	1
	Dock2	0	SHANK	1	0	1
	Dock2	0	SHANK3	1	0	1
	Dock2	0	EBP50	1	0	1
	Dock2	0	EBP50	2	0	2
	Dock2	0	KIAA0147	1	0	1
	Dock2	0	INADL	3	0	1
	Dock2	0	HEMBA 1003117	1	0	1
	Dock2	0	hAPXL	1	0	2
	Dock2	0	FLJ 10324	1	0	1
	Dock2	0	LIM-Mystique	1	0	1
	Dock2	0	LIM RIL	1	0	1
	Dock2	0	KIAA1634	5	0	1
	Dock2	0	BAI-1	6	0	1
	Dock2	0	Atrophin-1 Inter. Prot.	5	0	1
AA45	BLR-1	0	SHANK1	1	0	3
	BLR-1	0	SHANK3	1	0	3
	BLR-1	0	EBP50	1	0	3
	BLR-1	0	EBP50	2	0	3
	BLR-1	2	PDZK-1	2	5	1
AA56	Tax	0	TAX IP2	1	0	2
	Tax	0	Syntrophin gamma-2	1	0	1
	Tax	0	Syntrophin gamma-1	1	0	5
	Tax	0	KIAA0147	4	0	1
	Tax	0	KIAA0147	3	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	Tax	0	KIAA0147	2	0	5
	Tax	0	KIAA0147	1	0.1	5
	Tax	0	PTPL-1	2	0	2
	Tax	0	PTN-4	1	0	2
	Tax	0	INADL	3	0	1
	Tax	0	PSD95	3	0	1
	Tax	0	PSD95	2	0	1
	Tax	0	PSD95	1	0	5
	Tax	0	MUPP-1	13	0	5
	Tax	0	Outer Membrane	1	0	5
	Tax	0	NeDLG	3	1	5
	Tax	0	NeDLG	2	1	5
	Tax	0	FLJ 11215	1	0	1
	Tax	0	FLJ 10324	1	0	1
	Tax	0	FLJ 00011	1	0	1
	Tax	0	LIMK1	1	0	1
	Tax	0	LIM-Mystique	1	0	1
	Tax	0	Erbin	1	1	5
	Tax	0	LIM RIL	1	0	1
	Tax	0	DLG2	2	0	5
	Tax	0	DLG2	1	0	2
	Tax	0	DLG1	2	0	5
	Tax	0	DLG1	1	0.5	5
	Tax	0	Connector Enhancer	1	0	1
	Tax	0	KIAA1634	5	0	1
	Tax	0	BAI-1	6	0	1
	Tax	0	KIAA1634	4	0	2
	Tax	0	BAI-1	5	0	5
	Tax	0	KIAA1634	2	0	2
	Tax	0	KIAA1634	1	0.1	5
	Tax	0	BAI-1	4	0	2
	Tax	0	BAI-1	3	0	1
	Tax	0	BAI-1	2	0.5	5
	Tax	0	Atrophin-1 Inter. Prot.	5	0	3
	Tax	0	KIAA1526	1	0	3
	Tax	0	atrophin-1 interacting Protein	3	0	1
	Tax	0	atrophin-1 interacting Protein	2	0	1
	Tax	0	atrophin-1 interacting Protein	1	0	5
	Tax	0	AIPC	1	0	1
AA58	PAG	0	KIAA0382	1	0	1
	PAG	0	KIAA0316	1	0	1
	PAG	0	PIST	1	0	1
	PAG	0	hAPXL	1	0	2
	PAG	0	Outer Membrane	1	0	2
	PAG	0	SHANK	1	0	4
	PAG	0	SHANK3	1	0	2
	PAG	0	PDZK1	2,3,4	0	1
	PAG	0	FLJ 00011	1	0	3
	PAG	0	Atrophin-1 Inter. Prot.	5	0	1
AA59	PTEN	0	TIP1	1	0	2

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	PTEN	0	Syntrophin gamma-1	1	0	1
	PTEN	1.5	SHANK	1	5	3
	PTEN	0	INADL	8	0	1
	PTEN	0	PTPL-1	4	0	1
	PTEN	0.3	PTPL-1	2	1	4
	PTEN	0	PIST	1	0	1
	PTEN	0	HEMBA 1003117	1	0	1
	PTEN	0	MUPP-1	13	0	5
	PTEN	0	GRIP1	3	0	1
	PTEN	0	FLJ 10324	1	0	1
	PTEN	0	FLJ 00011	1	0	3
	PTEN	0	Mint 1	1,2	0	1
	PTEN	0	Mint 1	2	0	1
	PTEN	0	KIAA807		0	5
	PTEN	0	KIAA1634	2	0	5
	PTEN	0	BAI-1	3	0	2
	PTEN	0	Atrophin-1 Inter. Prot.	5	0	2
	PTEN	0	AIPC	1	0	1
	PTEN	0.3	KIAA0807(S)	1	0.5	5
AA60	AKT-1	2.5	TAX IP2	1	1	4
	AKT-1	0	KIAA807		0	1
	AKT-1	0	KIAA0807(S)	1	0	1
AA66.1	HPV E6 #66 (modified)	5	TIP1	1	1	5
	HPV E6 #66 (modified)	0	TAX IP2	1	0	2
	HPV E6 #66 (modified)	0	Syntrophin gamma-2	1	0	1
	HPV E6 #66 (modified)	0	Syntrophin gamma-1	1	0	1
	HPV E6 #66 (modified)	0	Synt. 1 alpha	1	0	2
	HPV E6 #66 (modified)	0	KIAA0147	1	0	2
	HPV E6 #66 (modified)	0	INADL	8	0	1
	HPV E6 #66 (modified)	0	PTPL-1	2	0	3
	HPV E6 #66 (modified)	0	PSD95	1,2,3	0	5
	HPV E6 #66 (modified)	0	PSD95	3	0	1
	HPV E6 #66 (modified)	0	PSD95	1	0	4
	HPV E6 #66 (modified)	0	MUPP-1	10	0	1
	HPV E6 #66 (modified)	0	MUPP-1	13	0	3
	HPV E6 #66 (modified)	1	NeDLG	1,2	0.5	5
	HPV E6 #66 (modified)	0	hAPXL	1	0	1
	HPV E6 #66 (modified)	0	Outer Membrane	1	0	5
	HPV E6 #66 (modified)	3.5	NeDLG	2	0.5	4
	HPV E6 #66 (modified)	0	NeDLG	1	0	1
	HPV E6 #66 (modified)	0	FLJ 10324	1	0	1
	HPV E6 #66 (modified)	0	FLJ 00011	1	0	1
	HPV E6 #66 (modified)	0	Mint 1	1,2	5	1
	HPV E6 #66 (modified)	0	Mint 1	2	0	1
	HPV E6 #66 (modified)	0	Erbin	1	0	1
	HPV E6 #66 (modified)	0	KIAA807		0	2
	HPV E6 #66 (modified)	0	DLG2	2	0	5
	HPV E6 #66 (modified)	0	DLG2	1	0	1
	HPV E6 #66 (modified)	0	DLG1	2	0	5
	HPV E6 #66 (modified)	0	DLG1	1	0	4
	HPV E6 #66 (modified)	5	DLG1	1,2	5	5
	HPV E6 #66 (modified)	0	BAI-1	5	5	1
	HPV E6 #66 (modified)	0	KIAA1634	2	0	1

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	HPV E6 #66 (modified)	0	KIAA1634	1	0	5
	HPV E6 #66 (modified)	0	BAI-1	3	5	1
	HPV E6 #66 (modified)	3	BAI-1	2	0.5	5
	HPV E6 #66 (modified)	0	Atrophin-1 Inter. Prot.	5	0	1
	HPV E6 #66 (modified)	0	KIAA1526	1	0	1
	HPV E6 #66 (modified)	0	atrophin-1 interacting Protein	1	0	5
	HPV E6 #66 (modified)	0	AIPC	1	0	1
	HPV E6 #66 (modified)	5	KIAA0807(S)	1	5	4
AA67.1	HPV E6 #57 (modified)	0	TIP1	1	0	0
	HPV E6 #57 (modified)	0	KIAA0147	1	0	1
	HPV E6 #57 (modified)	0	BAI-1	2	0	0
AA69.1	HPV E6 E16 (modified)	0	TIP1	1	0	3
	HPV E6 E16 (modified)	0	BAI-1	2	0	5
AA70.1	HPV E6 #16	0	TIP1	1	0	4
	HPV E6 #16	0	BAI-1	2	0	5
AA72.1	HPV E6 33 (modified)	0	ZO-2	1	5	1
	HPV E6 33 (modified)	0	TIP1	1	0	5
	HPV E6 33 (modified)	0	Syntrophin gamma-2	1	5	1
	HPV E6 33 (modified)	0	Synt. 1 alpha	1	1	3
	HPV E6 33 (modified)	0	SHANK	1	5	4
	HPV E6 33 (modified)	0	SHANK3	1	0	2
	HPV E6 33 (modified)	0	EBP50	1	0	2
	HPV E6 33 (modified)	0	EBP50	2	0	2
	HPV E6 33 (modified)	0	PTN-4	1	5	1
	HPV E6 33 (modified)	0	PSD95	1,2,3	0	5
	HPV E6 33 (modified)	5	PSD95	3	0.5	5
	HPV E6 33 (modified)	0	PSD95	1	5	2
	HPV E6 33 (modified)	0	PDZK1	2,3,4	5	1
	HPV E6 33 (modified)	0	Outer Membrane	1	0	5
	HPV E6 33 (modified)	0	NeDLG	3	5	1
	HPV E6 33 (modified)	0	NeDLG	2	5	2
	HPV E6 33 (modified)	0	NeDLG	1	5	1
	HPV E6 33 (modified)	0	NeDLG	1,2	0	5
	HPV E6 33 (modified)	0	MUPP-1	13	5	2
	HPV E6 33 (modified)	0	Mint 1	2	5	1
	HPV E6 33 (modified)	0	KIAA1634	1	0	5
	HPV E6 33 (modified)	0	KIAA1526	1	5	1
	HPV E6 33 (modified)	5	KIAA1095	1	0.5	5
	HPV E6 33 (modified)	0	KIAA0807(S)	1	0	5
	HPV E6 33 (modified)	0	KIAA0380	1	5	1
	HPV E6 33 (modified)	0	KIAA0316	1	5	2
	HPV E6 33 (modified)	0	KIAA0147	3	5	2
	HPV E6 33 (modified)	0	KIAA0147	1	0	5
	HPV E6 33 (modified)	0	hAPXL	1	1	3
	HPV E6 33 (modified)	0	FLJ 00011	1	5	1
	HPV E6 33 (modified)	0	DLG2	2	1	3
	HPV E6 33 (modified)	0	DLG2	1	5	1
	HPV E6 33 (modified)	5	DLG1	2	0.5	5
	HPV E6 33 (modified)	0	DLG1	1	1	3
	HPV E6 33 (modified)	0	BAI-1	6	5	1
	HPV E6 33 (modified)	0	BAI-1	5	5	1
	HPV E6 33 (modified)	0	BAI-1	2	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	HPV E8 33 (modified)	0	Atrophin-1 Inter. Prot.	5	5	1
	HPV E8 33 (modified)	5	Atrophin-1 Inter. Prot.	1	0.5	4
	HPV E8 33 (modified)	0	AIPC	1	5	1
AA74.1	HPV E6 52 (modified)	0	TIP1	1	0	0
	HPV E6 52 (modified)	0	BAI-1	2	0	5
AA75.1	HPV E6 58 (modified)	. 0	ZO-2	1	1	3
	HPV E6 58 (modified)	0	TIP1	1	0.5	4
	HPV E6 58 (modified)	0	Synt. 1 alpha	1	5	2
	HPV E6 58 (modified)	0	PSD95	1,2,3	0	5
	HPV E6 58 (modified)	0	PSD95	3	0	5
	HPV E6 58 (modified)	0	PSD95	1	0	5
	HPV E6 58 (modified)	0	PDZK1	2,3,4	5	1
	HPV E8 58 (modified)	5	Outer Membrane	1	0.5	5
	HPV E6 58 (modified)	5	NeDLG	3	5	2
	HPV E6 58 (modified)	0	NeDLG	2	0.5	5
	HPV E6 58 (modified)	0	NeDLG	1	5	1
	HPV E6 58 (modified)	0	NeDLG	1,2	0	5
	HPV E6 58 (modified)	0	MUPP-1	13	5	1
	HPV E6 58 (modified)	5	MUPP-1	10	3	3
	HPV E6 58 (modified)	0	Mint 1	2	5	1
	HPV E6 58 (modified)	0	KIAA1634	5	5	1
	HPV E6 58 (modified)	0	KIAA1634	2	5	1
	HPV E6 58 (modified)	0	KIAA1634	1	0	5
.	HPV E6 58 (modified)	0	KIAA1526	1	5	1
	HPV E6 58 (modified)	0	KIAA1095	1	5	1
	HPV E6 58 (modified)	0	KIAA0973	1	5	2
	HPV E6 58 (modified)	0	KIAA0807(S)	1	0	5
	HPV E6 58 (modified)	· 0	KIAA0380	1	5	1
	HPV E6 58 (modified)	0	KIAA0147	1	5	2
	HPV E6 58 (modified)	0	INADL	8	0.5	4
	HPV E6 58 (modified)	0	DLG2	2	0.5	5
	HPV E8 58 (modified)	0	DLG1	2	0	5
	HPV E8 58 (modified)	5	DLG1	1	0.5	5
	HPV E6 58 (modified)	0	BAI-1	5	5	2
	HPV E6 58 (modified)	0	BAI-1	4	5	2
	HPV E8 58 (modified)	0	BAI-1	3	5	2
	HPV E6 58 (modified)	0	BAI-1	2	0	5
	HPV E6 58 (modified)	0	Atrophin-1 Inter. Prot.	1	0	5
AA78.1	HPV E6 77 (Modified)	0	TIP1	1	0	0
	HPV E8 77 (Modified)	0	BAI-1	2	0	0
AA80.1	HPV E6 #35 (modified)	0	ZO-2	1	0	2
	HPV E6 #35 (modified)	0	ZO-1	1	0	1
	HPV E6 #35 (modified)	0	TIP1	1	0	5
	HPV E6 #35 (modified)	0	KIAA0382	1	0	2
	HPV E6 #35 (modified)	0	KIAA0380	1	0	3
	HPV E6 #35 (modified)	0	TAX IP2	1	0	4
	HPV E6 #35 (modified)	0	Syntrophin gamma-2	1	0	3
	HPV E6 #35 (modified)	0	Syntrophin gamma-1	1	0	4
	HPV E8 #35 (modified)	0	Synt. 1 alpha	1	0	5
	HPV E6 #35 (modified)	0	KIAA0147	4	0	1
	HPV E6 #35 (modified)	0.35	KIAA0147	3	5	4
	HPV E6 #35 (modified)	0	KIAA0147	2	0	5
	HPV E6 #35 (modified)	0	KIAA0147	1	0	5

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	HPV E6 #35 (modified)	0	INADL	8	0	4
	HPV E6 #35 (modified)	0	PTPL-1	4	0	1
	HPV E6 #35 (modified)	0	PTPL-1	2	0	2
	HPV E6 #35 (modified)	0	INADL	5	0	1
	HPV E6 #35 (modified)	0	PTN-4	1	0	4
	HPV E6 #35 (modified)	0	INADL	3	0	1
	HPV E6 #35 (modified)	0	PSD95	1,2,3	0	5
	HPV E6 #35 (modified)	0	PSD95	3	0	5
	HPV E6 #35 (modified)	0	PSD95	1	0	5
	HPV E6 #35 (modified)	0	PIST	1	0	1
	HPV E6 #35 (modified)	0	KIAA0973	1	0	2
	HPV E6 #35 (modified)	0	KIAA1095	1	0	4
	HPV E6 #35 (modified)	0	HEMBA 1003117	1	0	1
	HPV E6 #35 (modified)	0	MUPP-1	10	0	4
	HPV E6 #35 (modified)	0	MUPP-1	13	0	5
	HPV E6 #35 (modified)	0	NeDLG	1,2	0	5
	HPV E6 #35 (modified)	0	Outer Membrane	1	0	5
	HPV E6 #35 (modified)	0	NOS1	1	0	1
	HPV E6 #35 (modified)	0	NeDLG	3	0	5
	HPV E6 #35 (modified)	0	NeDLG	2	0	5
	HPV E6 #35 (modified)	0	NeDLG	1	0	5
	HPV E6 #35 (modified)	0	GRIP1	6	0	2
	HPV E6 #35 (modified)	0	GRIP1	3	0	2
	HPV E6 #35 (modified)	0	MUPP-1	5	0	2
	HPV E6 #35 (modified)	0	FLJ 12615 (PALS-1)	1	0	1
	HPV E6 #35 (modified)	0	FLJ 11215	1	0	4
	HPV E6 #35 (modified)	0	FLJ 10324	1	0	1
	HPV E6 #35 (modified)	0.35	FLJ 00011	1	5	3
	HPV E6 #35 (modified)	0	Mint 1	1,2	0	1
	HPV E6 #35 (modified)	0	Mint 1	2	0	2
	HPV E6 #35 (modified)	0	LIMK1	1	0	1
	HPV E6 #35 (modified)	0	LIM-Mystique	1	0	1
	HPV E6 #35 (modified)	0.4	Erbin	1	5	2
	HPV E6 #35 (modified)	0	LIM RIL	1	0	4
	HPV E6 #35 (modified)	0	KIAA807		0	5
	HPV E6 #35 (modified)	0.2	DLG2	2	0.5	5
	HPV E6 #35 (modified)	0	DLG2	1	0	5
	HPV E6 #35 (modified)	0	DLG1	3	5	3
	HPV E6 #35 (modified)	0	DLG1	2	0	5
	HPV E6 #35 (modified)	0	DLG1	1	0	5
	HPV E6 #35 (modified)	0	KIAA1719	5	0	1
	HPV E6 #35 (modified)	0	DLG1	1,2	0	5
	HPV E6 #35 (modified)	0	Connector Enhancer	1	0	1
	HPV E6 #35 (modified)	0	KIAA1634	5	0	3
	HPV E6 #35 (modified)	0	BAI-1	6	0	3
	HPV E6 #35 (modified)	0	KIAA1634	4	0	2
	HPV E6 #35 (modified)	0	BAI-1	5	0	5
	HPV E6 #35 (modified)	0	KIAA1634	2	0	3
	HPV E6 #35 (modified)	0	KIAA1634	1	0	5
	HPV E6 #35 (modified)	0	BAI-1	4	0	5
	HPV E6 #35 (modified)	0	BAI-1	3	0	4
	HPV E6 #35 (modified)	0	BAI-1	2	0	5
	HPV E6 #35 (modified)	0	Atrophin-1 Inter. Prot.	5	0	4

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
	HPV E6 #35 (modified)	1	KIAA1526	1	5	3
	HPV E6 #35 (modified)	0	atrophin-1 interacting Protein	3	0	4
	HPV E6 #35 (modified)	0	KIAA1284	1	0	1
	HPV E6 #35 (modified)	0.8	atrophin-1 interacting Protein	2	5	1
	HPV E6 #35 (modified)	0	atrophin-1 interacting Protein	1	0	5
	HPV E6 #35 (modified)	0	PDZ-73	2	0	2
	HPV E6 #35 (modified)	0	AIPC	1	5	1
	HPV E6 #35 (modified)	0.1	KIAA0807(S)	1	0.5	5
AA82	Adenovirus E4 Type9	0	ZO-2	1	0	3
	Adenovirus E4 Type9	0	ZO-1	1	0	2
	Adenovirus E4 Type9	0	KIAA0382	1	0	1
	Adenovirus E4 Type9	0	KIAA0300	1	0	1
	Adenovirus E4 Type9	0	INADL	8	0	2
	Adenovirus E4 Type9	0	PTPL-1	4	0	4
	Adenovirus E4 Type9	0.2	PTPL-1	2	5	3
	Adenovirus E4 Type9	0	PSD95	1,2,3	0	5
	Adenovirus E4 Type9	0.1	PSD95	1	5	4
	Adenovirus E4 Type9	0	PIST	1	0	1
	Adenovirus E4 Type9	0	KIAA1222	1	0	1
	Adenovirus E4 Type9	0.3	HEMBA 1003117	1	5	3
	Adenovirus E4 Type9	0.1	MUPP-1	11	5	5
	Adenovirus E4 Type9	0	NeDLG	1,2	0	5
	Adenovirus E4 Type9	0.1	Outer Membrane	1	5	5
	Adenovirus E4 Type9	0	NOS1	1	0	5
	Adenovirus E4 Type9	0.1	NeDLG	2	5	5
	Adenovirus E4 Type9	0	NeDLG	1	0	1
	Adenovirus E4 Type9	0	MUPP-1	10	0	1
	Adenovirus E4 Type9	0.1	FLJ 10324	1	5	3
	Adenovirus E4 Type9	0	FLJ 00011	1	0	1
	Adenovirus E4 Type9	0	Mint 1	1,2	0	2
	Adenovirus E4 Type9	0	Mint 1	2	0	2
	Adenovirus E4 Type9	0	KIAA807		0	4
	Adenovirus E4 Type9	0.05	DLG2	2	0.5	5
	Adenovirus E4 Type9	0.03	DLG1	2	0.3	5
	Adenovirus E4 Type9	0.1	DLG1	1	0.5	4
	Adenovirus E4 Type9	0	DLG1	1,2	0	5
	Adenovirus E4 Type9	0.1	Connector Enhancer	1	5	3
	Adenovirus E4 Type9	0	BAI-1	6	0	1
	Adenovirus E4 Type9	0.2	KIAA1634	4	5	4
	Adenovirus E4 Type9	0.15	KIAA1634	2	5	5
	Adenovirus E4 Type9	0.1	BAI-1	4	0.3	5
	Adenovirus E4 Type9	0.075	BAI-1	3	0.5	5
	Adenovirus E4 Type9	0	KIAA1634	1	0	5
	Adenovirus E4 Type9	0.02	BAI-1	2	0.3	5
	Adenovirus E4 Type9	0.1	atrophin-1 interacting Protein	3	5	4
	Adenovirus E4 Type9	0.02	atrophin-1 interacting Protein	1	0.5	5
	Adenovirus E4 Type9	0.2	KIAA0807(S)	1	5	3

AVC ID	AVC Name	Sequence	Accession No	GI
AA01.1	Clasp-1	VISKATPALPTVSISSSAEV		
AA02.1	Clasp-2	ISGTPTSTMVHGMTSSSSVV		
AA06	CD6	SPQPDSTDNDDYDDISAA	x60992	
AA07	CD34	QATSRNGHSARQHVVADTEL	m81104	
AA091	GAIP (G-alpha interacting protein) RGS 19	SSPTYRALLLQGPSQSSSEA	p49795 and X91809	1730188 and 1107697
AA092	alpha-1-syntrophin	IVFIIHSFLSAKVTRLGLLA	2209282A	1588880
AA093	neurofascin (chicken)	TEGNESSEATSPVNAIYSLA	CAA46330	63660
AA095	GluR5-2 (rat)	SFTSILTCHQRRTQRKETVA	M83561	204389
AA098L	ropporin	GPDGIITVNDFTQNPRVQLE	AAG27712	11037716
AA10	CD48	KKGTYLTDETHREVKFTSL	M58050	
AA105	CX43 (connexin 43)	PSSRASSRASSRPRPDDLEI	P17302	
AA106	Kir2.1 (inwardly rect. K+ channel)	LHNQASVPLEPRPLRRESEI	af153818S1 and AH009400	8132299
AA108.1	GLUR2 (glutamate receptor 2 -modified)	GGGGGSGGGGGSGIESVKI		
AA111	ephrin A2	RIAYSLLGLKDQVNTVGIPI	P29317 and XP_002088	125333 and 11427699
AA112	GluR delta-2	QPTPTLGLNLGNDPDRGTSI	AAC39579	2853315
AA113	SSTR2 (somatostatin recepor 2)	LNETTETQRTLLNGDLQTSI	XM_012697	12740782
AA114	GLUR7 (metabotropic glutamate receplor)	VDPNSPAAKKKYVSYNNLVI	XP_010942	12729188
AA115	presenilin-1	ATDYLVQPFMDQLAFHQFYI	XP_007441	11435042
AA116	MINT-2	KTMPAAMFRLLTGQETPLYI	AAC05308	2625029
AA117	presenilin-2	STDNLVRPFMDTLASHQLYI	NP 036618	7108360
AA118	MINT-1	KTMPAAMYRLLTAQEQPVYI	35430	6225060
AA121	CD68	ALVLIAFCIIRRRPSAYQAL	s57235	
AA123	a-actinin 2	VPGALDYAAFSSALYGESDL	p35609	543742
AA125	zona occludens 3 (ZO-3)	VHDAESSDEDGYDWGPATDL	NP_055243	10092691
AA13	CD96	KDITSDSENSNFRNEIQSLV		
AA140	KIA 1481	PIPAGGCTFSGIFPTLTSPL	AB040914	7959222
AA147	Na+/PI cotransporter 2	PPATPSPRLALPAHHNATRL	Q06495	730113
AA148L	CFTCR (cystic fibrosis transmembrane conductance regulator)	KPQIAALKEETEEEVQDTRL	AAC13867	306538
AA152L	ActRIIA	IVTVVTMVTNVDFPPKESSL	BAA08548	1321632
AA161	MINT-3	KTMPAATYRLLTGQEQPVYL	98018	6228953
AA169L	CAPON (carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase) mRNA	LLNVLQRQELGDGLDDEIAV	AF037070	2895554
AA172	RA-GEF (ras/rap1A-assoc.-GEF)	PYQSQGFSTEEDEDEQVSAV	NP_055062	7657261
AA177L	c-kit receptor	INSVGSTASSSQPLLVHDDV	TVHUKT	66811
AA178L	PDZ-binding kinase (PBK)	EDPKDRPSAAHIVEALETDV	XP_005110	11424184
AA180	NMDA Glutamate Receptor 2C (cysteine-free)	TQGFPGPATWRRISSLESEV		
AA182L	ephrin B2	ILNSIQVMRAQMNQIQSVEV	1F0MA	9256876
AA183L	RhoGAP 1 (PTPL1-associated)	PRLKRMQQFEDLEDEIPQFV	NP_004806 and NM_004815	4758882 and 4758881
AA185L	RGS12 (regulator of G-protein signaling 12	GPVPGEPAKPKTSAHHATFV	14924	3914623
AA190L	ephrin B1	PVYIVQEMPPQSPANIYYKV	XP_010388	11421689
AA192L	JAM (junctional adhesion molecule)	YSQPSARSEGEFKQTSSFLV	Q9Y624	10720061
AA205L	serotonin receptor 5-HT-2C	ENLELPVNPSSVVSERISSV	XP_013121	12743533
AA206L	CITRON protein	AGAVRTPLSQVNKVWDQSSV	O14578	6225217
AA207L	Nedasin (s-form)	RNIEEVYVGGKQVVPFSSSV	AAF13301	6469320
AA210L	APC- adenomatous polyposis coli protein	ESSGTQSPKRHSGSYLVTSV	P25054	114033
AA214L	ErbB-4 receptor	SLKPGTVLPPPPYRHRNTVV	q15303	3913590
AA215	CKR5 (HIV Co-receptor)	ERASSVYTRSTGEQEISVGL	P51681	
AA216	NMDA R2C	HPTDITGLPNLSDPSVSTVV	AAB59360	292283
AA217	catenin - delta 2	PYSELNYETSHYPASPDSWV	NP_001323	11034811
AA218	CSPG4 (chondroitin sulfae proteoglycan 4, melanoma-associated)	ELLQFCRTPNPALKNGQYWV	NM_001897 and X96753	4503098 and 1817313
AA22	DNAM-1	TREDIYVNYPTFSRRPKTRV		
AA220	claudin 10	GGEDFKTTNPSKQFDKNAYV	XP_007076	
AA222	claudin 18	DGGARTEDEVQSYPSKHDYV	XP_003116	
AA223	claudin 1	SYPTPRPYPKPAPSSGKDYV	XP_003151	
AA225	claudin 9	LGYSIPSRSGASGLDKRDYV	XP_012519	
AA226	claudin 7	KAGYRAPRSYPKSNSSKEYV	AAH01055	
AA227	claudin 2	PGQPPKVKSEFNSYSLTGYV	XP_010309	11420901
AA228	Nectin 2	SSPDSSYQGKGFVMSRAMYV	q92692	12643789
AA23.3	Fas Ligand	SSKSKSSEESQTFFGLYKL		
AA233L	serotonin receptor 5HT-2B	DTLLLTENEGDKTEEQVSYV	P41595	
AA240	Dopamine transporter	RELVDRGEVRQFTLRHWLKV	Q01959	266667

AVC ID	AVC Name	Sequence	Accession No	GI
AA243	alpha-2A Adrenergic receptor	HDFRRAFKKILARGDRKRIV	P08913	
AA244	alpha-28 Adrenergic receptor	QDFRRAFRRILARPWTQTAW	P18089	
AA245	alpha-2C Adrenergic receptor	DFRPSFKHILFRRARRGFRQ	P18825	
AA248	somatostatin receptor 4	EALQPEPGRKRIPLTRTTTF	P31391	
AA25	FceRIb	YSATYSELEDPGEMSPPIDL		
AA250	Serotonin receptor 3a	LAVLAYSITLVMLWSIWQYA	NP_000860	4504643
AA252	muscarinic Ach receptor M4	QQYQQRQSVIFHKRAPEQAL	P20309	
AA255	Clasp-5	RDSFHRSSFRKAETQLSQGS		
AA258	noradrenaline transporter	HHLVAQRDIRQFQLQHWLAI	M65015	189257
AA261	GA8A transporter 3	DAKLKSDGTIAAITEKETHF	XM_003161	12729857
AA262	glutamate transporter 3	NGGFAVDKSDTISFTQTSQF		11352332
AA264	bone morphogenelic protein receptor	TALRIKKTLAKMVESQDVKI	XM_015818	13648025
AA268	parathyroid hormone receptor 2	RPMESNPDTEGAQGETEDVL	P49190	
AA269	C5 Anaphylatoxin receptor	ESKSFTRSTVDTMAQKTQAV	P21730	
AA28.1	CDW125 (modified)	EVIGYIEKPGVETLEDSVF		
AA29.2	CDw1288	KDSRPSFVGSSSGHTSTTL		
AA29.3	IL-8RA	ARHRVTSYTSSSVNVSSNL		
AA30	LPAP	AWDDSARAAGGQGLHVTAL		
	LPAP	AAWDDSARAAGGQGLHVTAL		
AA300	TRAF2	NSYVRDDAIFIKAIVDLTGL	XM_011774	14737659
AA31	Mannose Receptor	GTSDMKDLVGNIEQNEHSVI		
AA38	Neuroligin	TFAAGFNSTGLPHSTTRV		
AA37	Glycophorin C	QGDPALQDAGDSSRKEYFI		
AA40	DOCK2	LASKSAEEGKQIPDSLSTDL		
AA45	BLR-1	PSWRRSSLSESENATSLTTF		
AA58	TAX	QISPGGLEPPSEKHFRETEV		
AA58	PAG	KENDYESISDLQQGRDITRL		
AA59	PTEN	DSDPENEPFDEDQHTQITKV		
AA60	AKT1	VDSERRPHFPQFSYSASSTA		
AA66.1	HPV E6 #66 (cysteine-free)	TGSALQAWRHTSRQATESTV		
AA67.1	HPV E6 #57 (cysteine-free)	HAMNAAPRAMENAPALRTSH		
AA69.1	HPV E6 #16 (Modified)	TGRGMSGGRSSRTRRETQL		
AA70.1	HPV E6 #18	SGGNRARQERLQRRRETQV		
AA72.1	HPV E8 33 (modified)	AAGGRSARGGRLQGRRETAL		
AA74.1	HPV E8 52 (modified)	SEGGRPTRGPRLQGRRVTQV		
AA75.1	HPV E8 58 (modified)	AVGGRPARGGRLQGRRQTQV		
AA78.1	HPV E6 77 (modified)	GGGRGSGLAGGSRGGGQSRQ		
AA80.1	HPV E6 #35 (cysteine-free)	GRWTGRAMSAWKPTRRETEV		
AA82	AdenoE4 typ9	VGTLLLERVIFPSVKIATLV		

Gene Name	GI	Domain Number	Sequence
26s subunit p27	9184389	1	RDMAEAHKEAMSRKLGQSESQGPPRAFAKVNSISPGSPSIAGLQVDDEIVEFGSVNTQNFQSLHNIGSVVQHSEGALAPTILLSVSM
AF6	430993	1	LRKEPEIITVTLKKQNGMGLSIVAAKGAGQDKLGIYVKSVVKGGAADVDGRLAAGDQLLSVDGRSLVGLSQERAAELMTRTSSVVTLEVAKQG
AIPC	12751451	1	LIRPSVISIIGLYKEKGKGLGFSIAGGRDCIRGQMGIFVKTIFPNGSAAEDGRLKEGDEILDVNGIPIKGLTFQEAIHTFKQIRSGLFVLTVRTKLVSPSLTNSS
AIPC	12751451	2	GISSLGRKTPGPKDRIVMEVTLNKEPRVGLGIGACCLALENSPPGIYIHSLAPGSVAKMESNLSRGDQILEVNSVNVRHAALSKVHAILSKCPPGPVRLVIGRHPNPKVSEQEMDEVIARSTYQESKEANSS
AIPC	12751451	3	QSENEEDVCFIVLNRKEGSGLGFSVAGGTDVEPKSITVHRVFSQGAASQEGTMNRGDFLLSVNGASLAGLAHGNVLKVLHQAQLHKDALVVIKKGMDQPRPS
AIPC	12751451	4	LGRSVAVHDALCVEVLKTSAGLGLSLDGGKSSVTGDGPLVIKRVYKGGAAEQAGIIEAGDEILAINGKPLVGLMHFDAWNIMKSVPEGPVQLLIRKHRNSS
alpha actinin-2 associated LIM protein	2773059	1	QTVILPGPAAWGFRLSGGIDFNQPLVITRITPGSKAAAANLCPGDVILAIDGFGTESMTHADGQDRIKAA
APXL-1	13651263	1	ILVEVQLSGGAPWGFTLKGGREHGEPLVITKIEEGSKAAAVDKLLAGDEIVGINDIGLSGFRQEAICLVKGSHKTLKLVVKRNSS
Atrophin-1 Interacting Protein	2947231	1	REKPLFTRDASQLKGTFLSTTLKKSNMGFGFTIIGGDEPDEFLQVKSVIPDGPAAQDGKMETGDVIVYINEVCVLGHTHADVVKLFQSVPIGQSVNLVLCRGY
Atrophin-1 Interacting Protein	2947231	2	LSGATQAELMTLTIVKGAQGFGFTIADSPTGQRVKQILDIQGCPGLCEGDLIVEINQQNVQNLSHTEVVDILKDCPIGSETSLIIHRGGFF
Atrophin-1 Interacting Protein	2947231	3	HYKELDVHLRRMESGFGFRILGGDEPGQPILIGAVIAMGSADRDGRLHPGDELVYVDGIPVAGKTHRYVIDLMHHAARNGQVNLTVRRKVLCG
Atrophin-1 Interacting Protein	2947231	4	EGRGISSHSLQTSDAVIHRKENEGFGFVIISSLNRPESGSTITVPHKIGRIIDGSPADRCAKLKVGDRILAVNGQSIINMPHADIVKLIKDAGLSVTLRIIPQEEL
Atrophin-1 Interacting Protein	2947231	5	LSDYRQPQDFDYFTVDMEKGAKGFGFSIRGGREYKMDLYVLRLAEDGPAIRNGRMRVGDQIIEINGESTRDMTHARAIELIKSGGRRVRLLLKRGTGQ
Atrophin-1 Interacting Protein	2947231	6	HESVIGRNPEGQLGFELKGGAENGQFPYLGEVKPGKVAYESGSKLVSEELLLEVNETPVAGLTIRDVLAVIKHCKDPLRLKCVKQGGIHR
BAI-1 Associated Protein	3370997	1	IQKKNHWTSRVHECTVKRGPQGELGVTVLGGAEHGEFPYVGAVAAVEAAGLPGGGEGPRLGEGELLLEVQGVRVSGLPRYDVLGVIDSCKEAVTFKAVRQ
BAI-1 Associated Protein	3370997	2	PSELKGKFIHTKLRKSSRGFGFTVVGGDEPDEFLQIKSLVLDGPAALDGKMETGDVIVSVNDTCVLGHTHAQVVKIFQSIPIGASVDLELCRGYPLPFDPDDP
BAI-1 Associated Protein	3370997	3	PATQPELITVHIVKGPMGFGFTIADSPGGGGQRVKQIVDSPRCRGLKEGDLIVEVNKKNVQALTHNQVVDMLVECPKGSEVTLLVQRGGNLS
BAI-1 Associated Protein	3370997	4	PDYQEQDIFLWRKETGFGFRILGGNEPGEPIYIGHIVPLGAADTDGRLRSGDELICVDGTPVIGKSHQLVVQLMQQAAKQGHVNLTVRRKVVFAVPKTENSS
BAI-1 Associated Protein	3370997	5	GVVSTVVQPYDVEIRRGENEGFGFVIVSSVSRPEAGTTFAGNACVAMPHKIGRIIEGSPADRCGKLKVGDRILAVNGCSITNKSHSDIVNLIKEAGNTVTLRIIP
BAI-1 Associated Protein	3370997	6	QATQEQDFYTVELERGAKGFGFSLRGGREYNMDLYVLRLAEDGPAERCGKMRIGDEILEINGETTKNMKHSRAIELIKNGGRRVRLFLKRG

Gene Name	GI	Domain Number	Sequence
CARD11	12382772	1	NLMFRKFSLERPFRPSVTSVGHVRGPGPSVQHTTLNGDSLTSQLTLLGGNARGSFVHSVKPGSLAEKAGLREGHQLLLLEGCIRGERQSVPLDTCTKEEAHWTIQRCSGPVTLHYKVNHEGYRKLV
CARD14	13129123	1	ILSQVTMLAFQGDALLEQISVIGGNLTGIFIHRVTPGSAADQMALRPGTQIVMVDYEASEPLFKAVLEDTTLEEAVGLLRRVDGFCCLSVKVNTDGYKRL
CASK	3087815	1	TRVRLVQFQKNTDEPMGITLKMNELNHCIVARIMHGGMIHRQGTLHVGDEIREINGISVANQTVEQLQKMLREMRGSITFKIVPSYRTQS
Connector Enhancer	3930780	1	LEQKAVLEQVQLDSPLGLEIHTTSNCQHFVSQVDTQVPTDSRLQIQPGDEVVQINEQVVVGWPRKNMVRELLREPAGLSLVLKKIPIP
Cytohesin Binding Protein	3192908	1	QRKLVTVEKQDNETFGFEIQSYRPQNQNACSSEMFTLICKIQEDSPAHCAGLQAGDVLANINGVSTEGFTYKQVVDLIRSSGNLLTIETLNG
DLG1	475816	1	IQVNGTDADYEYEEITLERGNSGLGFSIAGGTDNPHIGDDSSIFITKIITGGAAAQDGRLRVNDCILQVNEVDVRDVTHSKAVEALKEAGSIVRLYVKRRN
DLG1	475816	2	IQLIKGPKGLGFSIAGGVGNQHIPGDNSIYVTKIIEGGAAHKDGKLQIGDKLLAVNNVCLEEVTHEEAVTALKNTSDFVYLKVAKPTSMYMNDGN
DLG1	475816	3	ILHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDRIISVNSVDLRAASHEQAAAALKNAGQAVTIVAQYRPEEYSR
DLG2	12736552	1	ISYVNGTEIEYEFEEITLERGNSGLGFSIAGGTDNPHIGDDPGIFITKIIPGGAAAEDGRLRVNDCILRVNEVDVSEVSHSKAVEALKEAGSIVRLYVRRR
DLG2	12736552	2	ISVVEIKLFKGPKGLGFSIAGGVGNQHIPGDNSIYVTKIIDGGAAQKDGRLQVGDRLLMVNNYSLEEVTHEEAVAILKNTSEVVYLKVGNPTTI
DLG2	12736552	3	IWAVSLEGEPRKVVLHKGSTGLGFNIVGGEDGEGIFVSFILAGGPADLSGELQRGDQILSVNGIDLRGASHEQAAAALKGAGQTVTIIAQYQPED
DLG5	3650451	1	GIPYVEEPRHVKVQKGSEPLGISIVSGEKGGIYVSKVTVGSIAHQAGLEYGDQLLEFNGINLRSATEQQARLIIGQQCDTITILAQYNPHVHQLRNSSZLTD
DLG5	3650451	2	GILAGDANKKTLEPRVVFIKKSQLELGVHLCGGNLHGVFVAEVEDDSPAKGPDGLVPGDLILEYGSLDVRNKTVEEVYVEMLKPRDGVRLKVQYRPEEFIVTD
DVL1	2291005	1	LNIVTVTLNMERHHFLGISIVGQSNDRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSQTGPISLTVAKCW
DVL2	2291007	1	LNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDMNFENMSNDDAVRVLRDIVHKPGPIVLTVAKCWDPSPQNS
DVL3	6806886	1	IITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNEINFENMSNDDAVRVLREIVHKPGPITLTVAKCWDPSP
ELFIN 1	2957144	1	TTQQIDLQGPGPWGFRLVGRKDFEQPLAISRVTPGSKAALANLCIGDVITAIDGENTSNMTHLEAQNRIKGCTDNLTLTVARSEHKVWSPLV
ENIGMA	581636	1	IFMDSFKVVLEGPAPWGFRLQGGKDFNVPLSISRLTPGGKAAQAGVAVGDWVLSIDGENAGSLTHIEAQNKIRACGERLSLGLSRAQPV
ERBIN	8923908	1	QGHELAKQEIRVRVEKDPELGFSISGGVGGRGNPFRPDDDGIFVTRVQPEGPASKLLQPGDKIIQANGYSFINIEHGQAVSLLKTFQNTVELIIVREVSS

Gene Name	GI	Domain Number	Sequence
EZRIN Binding Protein 50	3220018	1	ILCCLEKGPNGYGFHLHGEKGKLGQYIRLVEPGSPAEKAGLLAGDRLVEVNGENVEKETHQQVVSRIRAALNAVRLLVVDPEFIVTD
EZRIN Binding Protein 50	3220018	2	IRLCTMKKGPSGYGFNLHSDKSKPGQFIRSVDPDSPAEASGLRAQDRIVEVNGVCMEGKQHGDVVSAIRAGGDETKLLVVDRETDEFFMNSS
FLJ00011	10440352	1	KNPSGELKTVTLSKMKQSLGISISGGIESKVQPMVKIEKIFPGGAAFLSGALQAGFELVAVDGENLEQVTHQRAVDTIRRAYRNKAREPMELVVRVPGPSPRP
FLJ11215	11436365	1	EGHSHPRVVELPKTEEGLGFNIMGGKEQNSPIYISRIIPGGIADRHGGLKRGDQLLSVNGVSVEGEHHEKAVELLKAAQGKVKLVVRYTPKVLEEME
FLJ12615	10434209	1	GQYGGETVKIVRIEKARDIPLGATVRNEMDSVIISRIVKGGAAEKSGLLHEGDEVLEINGIEIRGKDVNEVFDLLSDMHGTLTFVLIPSQQIKPPPA
FLJ20075	7019936	1	ILAHVKGIEKEVNVYKSEDSLGLTITDNGVGYAFIKRIKDGGVIDSVKTICVGDHIESINGENIVGWRHYDVAKKLKELKKEELFTMKLIEPKKAFEI
FLJ21687	10437836	1	KPSQASGHFSVELVRGYAGFGLTLGGGRDVAGDTPLAVRGLLKDGPAQRCGRLEVGDLVLHINGESTQGLTHAQAVERIRAGGPQLHLVIRRPLETHPGKP
GRIP 1	4539083	1	VVELMKKEGTTLGLTVSGGIDKDGKPRVSNLRQGGIAARSDQLDVGDYIKAVNGINLAKFRHDEIISLLKNVGERVVLEVEYE
GRIP 1	4539083	2	RSSVIFRTVEVTLHKEGNTFGFVIRGGAHDDRNKSRPVVITCVRPGGPADREGTIKPGDRLLSVDGIRLLGTTHAEAMSILKQCGQEAALLIEYDVSVMDSVAT
GRIP 1	4539083	3	HVATASGPLLVEVAKTPGASLGVALTTSMCCNKQVIVIDKIKSASIADRCGALHVGDHILSIDGTSMEYCTLAEATQFLANTTDQVKLEILPHHQTRLALKGPNSS
GRIP 1	4539083	4	TETTEVVLTADPVTGFGIQLQGSVFATETLSSPPLISYIEADSPAERCGVLQIGDRVMAINGIPTEDSTFEEASQLLRDSSITSKVTLEIEFDVAES
GRIP 1	4539083	5	AESVIPSSGTFHVKLPKKHNVELGITISSPSSRKPGDPLVISDIKKGSVAHRTGTLELGDKLLAIDNIRLDNCSMEDAVQILQQCEDLVKLKIRKDEDNSD
GRIP 1	4539063	6	IYTVELKRYGGPLGITISGTEEPFDPIIISSLTKGGLAERTGAIHIGDRILAINSSSLKGKPLSEAIHLLQMAGETVTLKIKKQTDAQSA
GRIP 1	4539083	7	IMSPTPVELHKVTLYKDSDMEDFGFSVADGLLEKGVYVKNIRPAGPGDLGGLKPYDRLLQVNHVRTRDFDCCLVVPLIAESGNKLDLVISRNPLA
GTPase Activating Enzyme	2389008	1	LALPRDGQGRLGFEVDAEGFVTHVERFTFAETAGLRPGARLLRVCGQTLPSLRPEAAAQLLRSAPKVC
Guanine Exchange Factor	6650765	1	AKAKWRQVVLQKASRESPLQFSLNGGSEKGFGIFVEGVEPGSKAADSGLKRGDQIMEVNGQNFENITFMKAVEILRNNTHLALTVKTNIFVFKEL
HEMBA 1000505	10436387	1	LENVIAKSLLIKSNEGSYGFGLEDKNKVPIIKLVEKGSNAEMAGMEVGKKIFAINGDLVFMRPFNEVDCFLKSCLNSRKPLRVLVSTKP
HEMBA 1000505	10436367	2	PRETVKIPDSADGLGFQIRGFGPSVVHAVGRGTVAAAAGLHPGQCIIKVNGINVSKETHASVIAHVTACRKYRRPTKQDSIQ
HEMBA 1003117	7022001	1	EDFCYVFTVELERGPSGLGMGLIDGMHTHLGAPGLYIQTLLPGSPAAADGRLSLGDRILEVNGSSLLGLGYLRAVDLIRHGGKKMRFLVAKSDVETAKKI
INADL	2370148	1	IWQIEYIDIERPSTGGLGFSVVALRSQNLGKVDIFVKDVQPGSVADRDQRLKENDQILAINHTPLDQNISHQQAIALLQQTTGSLRLIVAREPVHTKSSTSSSE

Gene Name	GI	Domain Number	Sequence
INADL	2370148	2	PGHVEEVELINDGSGLGFGIVGGKTSGVVVRTIVPGGLADRDGRLQTGDHILKIGGTNVQGMTSEQVAQVLRNCGNSS
INADL	2370148	3	PGSDSSLFETYNVELVRKDGQSLGIRIVGYVGTSHTGEASGIYVKSIIPGSAAYHNGHIQVNDKIVAVDGVNIQGFANHDVVEVLRNAGQVVHLTLVRRKTSS
INADL	2370148	4	NSDDAELQKYSKLLPIHTLRLGVEVDSFDGHHYISSIVSGGPVDTLGLLQPEDELLEVNGMQLYGKSRREAVSFLKEVPPPFTLVCCRRLFDDEAS
INADL	2370148	5	LSSPEVKIVELVKDCKGLGFSILDYQDPLDPTRSVIVIRSLVADGVAERSGGLLPGDRLVSVNEYCLDNTSLAEAVEILKAVPPGLVHLGICKPLVEFIVTD
INADL	2370148	6	PNFSHWGPPRIVEIFREPNVSLGISIVVGQTVIKRLKNGEELKGIFIKQVLEDSPAGKTNALKTGDKILEVSGVDLQNASHSEAVEAIKNAGNPVVFIVQSLSSTPRVIPNVHNKANSS
INADL	2370148	7	PGELHIIELEKDKNGLGLSLAGNKDRSRMSIFVVGINPEGPAAADGRMRIGDELLEINNQILYGRSHQNASAIIKTAPSKVKLVFIRNEDAVNQMANSS
INADL	2370148	8	PATCPIVPGQEMIIEISKGRSGLGLSIVGGKDTPLNAIVIHEVYEEGAAARDGRLWAGDQILEVNGVDLRNSSHEEAITALRQTPQKVRLVVY
KIAA0147	1469875	1	ILTLTILRQTGGLGISIAGGKGSTPYKGDDEGIFISRVSEEGPAARAGVRVGDKLLEVNGVALQGAEHHEAVEALRGAGTAVQMRVWRERMVEPENAEFIVTD
KIAA0147	1469875	2	PLRQRHVACLARSERGLGFSIAGGKGSTPYRAGDAGIFVSRIAEGGAAHRAGTLQVGDRVLSINGVDVTEARHDHAVSLLTAASPTIALLLEREAGG
KIAA0147	1469875	3	ILEGPYPVEEIRLPRAGGPLGLSIVGGSDHSSHPFGVQEPGVFISKVLPRGLAARSGLRVGDRILAVNGQDVRDATHQEAVSALLRPCLELSLLVRRDPAEFIV
KIAA0147	1469875	4	RELCIQKAPGERLGISIRGGARGHAGNPRDPTDEGIFISKVSPTGAAGRDGRLRVGLRLLEVNQQSLLGLTHGEAVQLLRSVGDTLTVLVCDGFEASTDAALE
KIAA0303	2224546	1	PHQPIVIHSSGKNYGFTIRAIRVYVGDSDIYTVHHIVWNVEEGSPACQAGLKAGDLITHINGEPVHGLVHTEVIELLLKSGNKVSITTTPF
KIAA0313	7657260	1	ILACAAKAKRRLMTLTKPSREAPLPFILLGGSEKGFGIFVDSVDSGSKATEAGLKRGDQILEVNGQNFENIQLSKAMEILRNNTHLSITVKTNLFVFKELLTNSS
KIAA0316	6683123	1	IPPAPRKVEMRRDPVLGFGFVAGSEKPVVVRSVTPGGPSEGKLIPGDQIVMINDEPVSAAPRERVIDLVRSCKESILLTVIQPYPSPK
KIAA0340	2224620	1	LNKRTTMPKDSGALLGLKVVGGKMTDLGRLGAFITKVKKGSLADVVGHLRAGDEVLEWNGKPLPGATNEEVYNIILESKSEPQVEIIVSRPIGDIPRIHRD
KIAA0380	2224700	1	QRCVIIQKDQHGFGFTVSGDRIVLVQSVRPGGAAMKAGVKEGDRIIKVNGTMVTNSSHLEVVKLIKSGAYVALTLLGSS
KIAA0382	7662087	1	ILVQRCVIIQKDDNGFGLTVSGDNPVFVQSVKEDGAAMRAGVQTGDRIIKVNGTLVTHSNHLEVVKLIKSGSYVALTVQGRPPGNSS
KIAA0440	2662160	1	SVEMTLRRNGLGQLGFHVNYEGIVADVEPYGYAWQAGLRQGSRLVEICKVAVATLSHEQMIDLLRTSVTVKVVIIPPHD
KIAA0545	14762850	1	LKVMTSGWETVDMTLRRNGLGQLGFHVKYDGTVAEVEDYGFAWQAGLRQGSRLVEICKVAVVTLTHDQMIDLLRTSVTVKVVIIPPFEDGTPRRGW

Gene Name	GI	Domain Number	Sequence
KIAA0559	3043641	1	HYIFPHARIKITRDSKDHTVSGNGLGIRIVGGKEIPGHSGEIGAYIAKILPGGSAEQTGKLMEGMQVLEWNGIPLTSKTYEEVQSIISQQSGEAEICVRLDLNML
KIAA0561	3043845	1	LCGSLRPPIVIHSSGKKYGFSLRAIRVYMGDSDVYTVHHVVWSVEDGSPAQEAGLRAGDLITHINGESVLGLVHMDVVELLLKSGNKISLRTTALENTSIKVG
KIAA0613	3327039	1	SYSVTLTGPGPWGFRLQGGKDFNMPLTISRITPGSKAAQSQLSQGDLVVAIDGVNTDTMTHLEAQNKIKSASYNLSLTLQKSKNSS
KIAA0751	12734165	1	ISRDSGAMLGLKVVGGKMTESGRLCAFITKVKKGSLADTVGHLRPGDEVLEWNGRLLQGATFEEVYNIILESKPEPQVELVVSRPIAIHRD
KIAA0807	3882334	1	ISALGSMRPPIIIHRAGKKYGFTLRAIRVYMGDSDVYTVHHMVWHVEDGGPASEAGLRQGDLITHVNGEPVHGLVHTEVVELILKSGNKVAISTTPLENSS
KIAA0858	4240204	1	FSDMRISINQTPGKSLDFGFTIKWDIPGIFVASVEAGSPAEFSQLQVDDEIIAINNTKFSYNDSKEWEEAMAKAQETGHLVMDVRRYGKAGSPE
KIAA0902	4240292	1	QSAHLEVIQLANIKPSEGLGMYIKSTYDGLHVITGTTENSPADRCKKIHAGDEVIQVNHQTVVGWQLKNLVNALREDPSGVILTLKKRPQSMLTSAPA
KIAA0967	4589577	1	ILTQTLIPVRHTVKIDKDTLLQDYGFHISESLPLTVVAVTAGGSAHGKLFPGDQILQMNNEPAEDLSWERAVDILREAEDSLSITVVRCTSGVPKSSNSS
KIAA0973	4589589	1	GLRSPITIQRSGKKYGFTLRAIRVYMGDTDVYSVHHIVWHVEEGGPAQEAGLCAGDLITHVNGEPVHGMVHPEVVELILKSGNKVAVTTTPFE
KIAA1095	5889526	1	QGEETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDGSSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVNGRDLSRATHDQAVEAFKTAKEPIVVQVLRRTPRTKM
KIAA1095	5889526	2	QEMDREELELEEVDLYRMNSQDKLGLTVCYRTDDEDDIGIYISEIDPNSIAAKDGRIREGDRIIQINGIEVQNREEAVALLTSEENKNFSLLIARPELQLD
KIAA1202	6330421	1	RSFQYVPVQLQGGAPWGFTLKGGLEHCEPLTVSKIEDGGKAALSQKMRTGDELVNINGTPLYGSRQEALILIKGSFRILKLIVRRRNAPVS
KIAA1222	6330610	1	ILEKLELFPVELEKDEDGLGISIIGMGVGADAGLEKLGIFVKTVTEGGAAQRDGRIQVNDQIVEVDGISLVGVTQNFAATVLRNTKGNVRFVIGREKPGQVS
KIAA1284	6331369	1	KDVNVYVNPKKLTVIKAKEQLKLLEVLVGIIHQTKWSWRRTGKQGDGERLVVHGLLPGGSAMKSGQVLIGDVLVAVNDVDVTTENIERVLSCIPGPMQVKLTFENAYDVKRET
KIAA1389	7243158	1	TRGCETVEMTLRRNGLGQLGFHVNFEGIVADVEPFGFAWKAGLRQGSRLVEICKVAVATLTHEQMIDLLRTSVTVKVVIIQPHDDGSPRR
KIAA1415	7243210	1	VENILAKRLLILPQEEDYGFDIEEKNKAVVVKSVQRGSLAEVAGLQVGRKIYSINEDLVFLRPFSEVESILNQSFCSRRPLRLLVATKAKEIIKIP
KIAA1525	5817188	1	PDSAGPGEVRLVSLRRAKAHEGLGFSIRGGSEHGVGIYVSLVEPGSLAEKEGLRVGDQILRVNDKSLARVTHAEAVKALKGSKKLVLSVYSAGRIPGGYVTN
KIAA1526	5817166	2	LQGGDEKKVNLVLGDGRSLGLTIRGGAEYGLGIYITGVDPGSEAEGSGLKVGDQILEVNWRSFLNILHDEAVRLLKSSRHLILTVKDVGRLPHARTTVDE
KIAA1526	5817166	3	WTSGAHVHSGPCEEKCGHPGHRQPLPRIVTIQRGGSAHNCGQLKVGHVILEVNGLTLRGKEHREAARIIAEAFKTKDRDYIDFLDSL

Gene Name	GI	Domain Number	Sequence
KIAA1620	10047316	1	ELRRAELVEIIVETEAQTGVSGINVAGGGKEGIFVRELREDSPAARSLSLQEGDQLLSARVFFENFKYEDALRLLQCAEPYKVSFCLKRTVPTGDLALRP
KIAA1634	10047344	1	PSQLKGVLVRASLKKSTMGFGFTIIGGDRPDEFLQVKNVLKDGPAAQDGKIAPGDVIVDINGNCVLGHTHADVVQMFQLVPVNQYVNLTLCRGYPLPDDSED
KIAA1634	10047344	2	ASSGSSQPELVTIPLIKGPKGFGFAIADSPTGQKVKMILDSQWCQGLQKGDIIKEIYHQNVQNLTHLQVVEVLKQFPVGADVPLLILRGGPPSPTKTAKM
KIAA1634	10047344	3	LYEDKPPLTNTFLISNPRTTADPRILYEDKPPNTKDLDVFLRKQESGFGFRVLGGDGPDQSIYIGAIIPLGAAEKDGRLRAADELMCIDGIPVKGKSHKQVLDLMTTAARNGHVLLTVRRKIFYGEKQPEDDSGSPGIH
KIAA1634	10047344	4	PAPQEPYDVVLQRKENEGFGFVILTSKNKPPPGVIPHKIGRVIEGSPADRCGKLKVGDHISAVNGQSIVELSHDNIVQLIKDAGVTVTLTVIAEEEHHGPPS
KIAA1634	10047344	5	QNLGCYPVELERGPRGFGFSLRGGKEYNMGLFILRLAEDGPAIKDGRIHVGDQIVEINGEPTQGITHTRAIELIQAGGNKVLLLLRPGTGLIPDHGLA
KIAA1719	1267982	0	ITVVELIKKEGSTLGLTISGGTDKDGKPRVSNLRPGGLAARSDLLNIGDYIRSVNGIHLTRLRHDEIITLLKNVGERVVLEVEY
KIAA1719	1267982	1	ILDVSLYKEGNSFGFVLRGGAHEDGHKSRPLVLTYVRPGGPADREGSLKVGDRLLSVDGIPLHGASHATALATLRQCSHEALFQVEYDVATP
KIAA1719	1267982	2	IHTVANASGPLMVEIVKTPGSALGISLTTTSLRNKSVITIDRIKPASVVDRSGALHPGDHILSIDGTSMEHCSLLEATKLLASISEKVRLEILPVPQSQRPL
KIAA1719	1267982	3	IQIVHTETTEVVLCGDPLSGFGLQLQGGIFATETLSSPPLVCFIEPDSPAERCGLLQVGDRVLSINGIATEDGTMEEANQLLRDAALAHKVVLEVEFDVAESV
KIAA1719	1267982	4	IQFDVAESVIPSSGTFHVKLPKKRSVELGITISSASRKRGEPLIISDIKKGSVAHRTGTLEPGDKLLAIDNIRLDNCPMEDAVQILRQCEDLVKLKIRKDEDN
KIAA1719	1267982	5	IQTTGAVSYTVELKRYGGPLGITISGTEEPFDPIVISGLTKRGLAERTGAIHVGDRILAINNVSLKGRPLSEAIHLLQVAGETVTLKIKKQLDR
KIAA1719	1267982	6	ILEMEELLLPTPLEMHKVTLHKDPMRHDFGFSVSDGLLEKGVYVHTVRPDGPAHRGGLQPFDRVLQVNHVRTRDFDCCLAVPLLAEAGDVLELIISRKPHTA
LIM Mystique	12734250	1	MALTVDVAGPAPWGFRITGGRDFHTPIMVTKVAERGKAKDADLRPGDIIVAINGESAEGMLHAEAQSKIRQSPSPLRLQLDRSQATSPGQT
LIM Protein	3108092	1	SNYSVSLVGPAPWGFRLQGGKDFNMPLTISSLKDGGKAAQANVRIGDVVLSIDGINAQGMTHLEAQNKIKGCTGSLNMTLQRAS
LIM-RIL	1085021	1	IHSVTLRGPSPWGFRLVGRDFSAPLTISRVHAGSKASLAALCPGDLIQAINGESTELMTHLEAQNRIKGCHDHLTLSVSRPE
LIMK1	4587498	1	TLVEHSKLYCGHCYYQTVVTPVIEQILPDSPGSHLPHTVTLVSIPASSHGKRGLSVSIDPPHGPPGCGTEHSHTVRVQGVDPGCMSPDVKNSIHVGDRILEINGTPIRNVPLDEIDLLIQETSRLLQLTLEHD
LIMK2	1805593	1	PYSVTLISMPATTEGRRGFSVSVESACSNYATTVQVKEVNRMHISPNNRNAIHPGDRILEINGTPVRTLRVEEVEDAISQTSQTLQLLIEHD
LU-1	U52111 (acc. #)	1	VCYRTDDEEDLGIYVGEVNPNSIAAKDGRIREGDRIIQINGVDVQNREEAVAILSQEENTNISLLVARPE

Gene Name	GI	Domain Number	Sequence
MINT1	2625024	1	SENCKdVFIEKQKGEILGVVIVESGWGSILPTVIIANMMHGGPAEKSGKLNIGDQIMSINGTSLVGLPLSTCQSIIKGLKNQSRVKLNIVRCPPVNSS
MINT1	2625024	2	LRCPPVTTVLIRRPDLRYQLGFSVQNGIICSLMRGGIAERGGVRVGHRIIEINGQSVVATPHEKIVHILSNAVGEIHMKTMPAAMYRLLNSS
MINT3	3169808	1	LSNSDNCREVHLEKRRGEGLGVALVESGWGSLLPTAVIANLLHGGPAERSGALSIGDRLTAINGTSLVGLPLAACQAAVRETKSQTSVTLSIVHCPPVTTAIM
MINT3	3169808	2	LVHCPPVTTAIIHRPHAREQLGFCVEDGIICSLLRGGIAERGGIRVGHRIIEINGQSVVATPHARIIELLTEAYGEVHIKTMPAATYRLLTG
MPP1	189785	1	RKVRLIQFEKVTEEPMGITLKLNEKQSCTVARILHGGMIHRQGSLHVGDEILEINGTNVTNHSVDQLQKAMKETKGMISLKVIPNQ
MPP2	939884	1	PVPPDAVRMVGIRKTAGEHLGVTFRVEGGELVIARILHGGMVAQQGLLHVGDIIKEVNGQPVGSDPRALQELLRNASGSVILKILPNYQ
MUPP1	2104784	1	QGRHVEVFELLKPPSGGLGFSVVGLRSENRGELGIFVQEIQEGSVAHRDGRLKETDQILAINGQALDQTITHQQAISILQKAKDTVQLVIARGSLPQLV
MUPP1	2104784	2	PVHWQHMETIELVNDGSGLGFGIIGGKATGVIVKTILPGGVADQHGRLCSGDHILKIGDTDLAGMSSEQVAQVLRQCGNRVKLMIARGAIEERTAPT
MUPP1	2104784	3	QESETFDVELTKNVQGLGITIAGYIGDKKLEPSGIFVKSITKSSAVEHDGRIQIGDQIIAVDGTNLQGFTNQQAVEVLRHTGQTVLLTLMRRGMKQEA
MUPP1	2104784	4	LNYEIVVAHVSKFSENSGLGISLEATVGHHFIRSVLPEGPVGHSGKLFSGDELLEVNGITLLGENHQDVVNILKELPIEVTMVCCRRTVPPT
MUPP1	2104784	5	WEAGIQHIELEKGSKGLGFSILDYQDPIDPASTVIIRSLVPGGIAEKDGRLLPGDRLMFVNDVNLENSSLEEAVEALKGAPSGTVRIGVAKPLPLSPEE
MUPP1	2104784	6	RNVSKESFERTINIAKGNSSLGMTVSANKDGLGMIVRSIIHGGAISRDGRIAIGDCILSINEESTISVTNAQARAMLRRHSLIGPDIKITYVPAEHLEE
MUPP1	2104784	7	LNWNQPRRVELWREPSKSLGISIVGGRGMGSRLSNGEVMRGIFIKHVLEDSPAGKNGTLKPGDRIVEVDGMDLRDASHEQAVEAIRKAGNPVVFMVQSIINR
MUPP1	2104784	8	LTGELHMIELEKGHSGLGLSLAGNKDRSRMSVFIVGIDPNGAAGKDGRLQIADELLEINGQILYGRSHQNASSIIKCAPSKVKIIFIRNKDAVNQ
MUPP1	2104784	9	LSSFKNVQHLELPKDQGGLGIAISEEDTLSGVIIKSLTEHGVAATDGRLKVGDQILAVDDEIVVGYPIEKFISLLKTAKMTVKLTIHAENPDSQ
MUPP1	2104784	10	LPGCETTIEISKGRTGLGLSIVGGSDTLLGAIIIHEVYEEGAACKDGRLWAGDQILEVNGIDLRKATHDEAINVLRQTPDRVRLTLYRDEAPYKE
MUPP1	2104784	11	KEEEVCDTLTIELQKKPGKGLGLSIVGKRNDTGVFVSDIVKGGIADADGRLMQGDQILMVNGEDVRNATQEAVAALLKCSLGTVTLEVGRIKAGPFHS
MUPP1	2104784	12	LQGLRTVEMKKGPTDSLGISIAGGVGSPLGDVPIFIAMMHPTGVAAQTQKLRVGDRIVTICGTSTEGMTHTQAVNLLKNASGSIEMQVVAGGDVSV
MUPP1	2104784	13	LGPPQCKSITLERGPDGLGFSIVGGYGSPHGDLPIYVKTVFAKGAASEDGRLKRGDQIIAVNGQSLEGVTHEEAVAILKRTKGTVTLMVLS

Gene Name	GI	Domain Number	Sequence
NeDLG	10863920	1	IQYEEIVLERGNSGLGFSIAGGIDNPHVPDDPGIFITKIIPGGAAAMDGRLGVNDCVLRVNEVEVSEVVHSRAVEALKEAGPVVRLVVRRRQN
NeDLG	10863920	2	ITLLKGPKGLGFSIAGGIGNQHIPGDNSIYITKIIEGGAAQKDGRLQIGDRLLAVNNTNLQDVRHEEAVASLKNTSDMVYLKVAKPGSLE
NeDLG	10863920	3	ILLHKGSTGLGFNIVGGEDGEGIFVSFILAGGPADLSGELRRGDRILSVNGVNLRNATHEQAAAALKRAGQSVTIVAQYRPEEYSRFESKIHDLREQMMNSSMSSGSGSLRTSEKRSLE
NOS1	642525	1	IQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQAGDIILAVNGRPLVDLSYDSALEVLRGIASETHVVLILRGP
novel PDZ gene	7228177	1	QANSDESDIIHSVRVEKSPAGRLGFSVRGGSEHGLGIFVSKVEEGSSAERAGLCVGDKITEVNGLSLESTTMGSAVKVLTSSSRLHMMVRRMGRVPGIKFSK
novel PDZ gene	7228177	2	PSDTSSEDGVRRIVHLYTTSDDFCLGFNIRGGKEFGLGIYVSKVDHGGLAEENGIKVGDQVLAANGVRFDDISHSQAVEVLKGQTHIMLTIKETGRYPAYKEMN
Novel Serine Protease	1821243	1	KIKKFLTESHDRQAKGKAITKKKYIGIRMMSLTSSKAKELKDRHRDFPDVISGAYIIEVIPDTPAEAGGLKENDVIISINGQSVVSANDVSDVIKRESTLNMVVRRG
Outer Membrane	7023825	1	LLTEEEINLTRGPSGLGFNIVGGTDQQYVSNDSGIYVSRIKENGAAALDGRLQEGDKILSVNGQDLKNLLHQDAVDLFRNAGYAVSLRVQHRLQVQNGIHS
p55T	12733367	1	PVDAIRILGIHKRAGEPLGVTFRVENNDLVIARILHGGMIDRQGLLHVGDIIKEVNGHEVGNNPKELQELLKNISGSVTLKILPSYRDTITPQQ
PAR3	8037914	1	DDMVKLVEVPNDGGPLGIHVVPFSARGGRTLGLLVKRLEKGGKAEHENLFRENDCIVRINDGDLRNRRFEQAQHMFRQAMRTPIIWFHVVPAA
PAR3	8037914	2	GKRLNIQLKKGTEGLGFSITSRDVTIGGSAPIYVKNILPRGAAIQDGRLKAGDRLIEVNGVDLVGKSQEEVVSLLRSTKMEGTVSLLVFRQEDA
PAR3	8037914	3	TPDGTREFLTFEVPLNDSGSAGLGVSVKGNRSKENHADLGIFVKSIINGGAASKDGRLRVNDQLIAVNGESLLGKTNQDAMETLRRSMSTEGNKRGMIQLIVA
PAR6	2613011	1	LPETHRRVRLHKHGSDRPLGFYIRDGMSVRVAPQGLERVPGIFISRLVRGGLAESTGLLAVSDEILEVNGIEVAGKTLDQVTDMMVANSHNLIVTVKPANQR
PAR6 GAMMA	13537118	1	IDVDLVPETHRRVRLHRHGCEKPLGFYIRDGASVRVTPHGLEKVPGIFISRMVPGGLAESTGLLAVNDEVLEVNGIEVAGKTLDQVTDMMIANSHNLIVTVKPA
PDZ-73	5031978	1	RSRKLKEVRLDRLHPEGLGLSVRGGLEFGCGLFISHLIKGGQADSVGLQVGDEIVRINGYSISSCTHEEVINLIRTKKTVSIKVRHIGLIPVKSSPDEFH
PDZ-73	5031978	2	IPGNRENKEKKVFISLVGSRGLGCSISSGPIQKPGIFISHVKPGSLSAEVGLEIGDQIVEVNGVDFSNLDHKEAVNVLKSSRSLTISIVAAAGRELFMTDEF
PDZ-73	5031978	3	PEQIMGKDVRLLRIKKEGSLDLALEGGVDSPIGKVVVSAVYERGAAERHGGIVKGDEIMAINGKIVTDYTLAEADAALQKAWNQGGDWIDLVVAVCPPKEYDD
PDZK1	2944188	1	LTSTFNPRECKLSKQEGQNYGFFLRIEKDTEGHLVRVVEKCSPAEKAGLQDGDRVLRINGVFVDKEEHMQVVDLVRKSGNSVTLLVLDGDSYEKAGSPGIHR
PDZK1	2944188	2	RLCYLVKEGGSYGFSLKTVQGKKGVYMTDITPQGVAMRAGVLADDHLIEVNGENVEDASHEEVVEKVKKSGSRVMFLLVDKETDKREFIVTD

Gene Name	GI	Domain Number	Sequence
PDZK1	2944186	3	QFKRETASLKLLPHQPRIVEMKKGSNGYGFYLRAGSEQKGQIIKDIDSGSPAEEAGLKNNDLVVAVNGESVETLDHDSVVEMIRKGGDQTSLLVVDKETDNM
PDZK1	2944188	4	PDTTEEVDHKPKLCRLAKGENGYGFHLNAIRGLPGSFIKEVQKGGPADLAGLEDEDVIIEVNGVNVLDEPYEKVVDRIQSSGKNVTLLVZGKNSS
PICK1	4678411	1	PTVPGKVTLQKDAQNLIGISIGGGAQYCPCLYIVQVFDNTPAALDGTVAAGDEITGVNGRSIKGKTKVEVAKMIQEVKGEVTIHYNKLQ
PIST	98374330	1	SQGVGPIRKVLLLKEDHEGLGISITGGKEHGVPILISEIHPGQPADRCGGLHVGDAILAVNGVNLRDTKHKEAVTILSQQRGEIEFEVVYVAPEVDSD
prIL16	1478492	1	IHVTILHKEEGAGLGFSLAGGADLENKVITVHRVFPNGLASQEGTIQKGNEVLSINGKSLKGTTHHDALAILRQAREPRQAVIVTRKLTPEEFIVTD
prIL16	1478492	2	TAEATVCTVTLEKMSAGLGFSLEGGKGSLHGDKPLTINRIFKGAASEQSETVQPGDEILQLGGTAMQGLTRFEAWNIIKALPDGPVTIVIRRKSLQSK
PSD95	3318652	1	LEYEeITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIPGGAAAQDGRLRVNDSILFVNEVDVREVTHSAAVEALKEAGSIVRLYVMRRKPPAENSS
PSD95	3318652	2	HVMRRKPPAEKVMEIKLIKGPKGLGFSIAGGVGNQHIPGDNSIYVTKIIEGGAAHKDGRLQIGDKILAVNSVGLEDVMHEDAVAALKNTYDVVYLKVAKPSNAY
PSD95	3318652	3	REDIPREPRRIVIHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDQILSVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEFIVTD
PTN-3	179912	1	LIRITPDEDGKFGFNLKGGVDQKMPLVVSRINPESPADTCIPKLNEGDQIVLINGRDISEHTHDQVVMFIKASRESHSRELALVIRRR
PTN-4	190747	1	IRMKPDENGRFGFNVKGGYDQKMPVIVSRVAPGTPADLCVPRLNEGDQVVLINGRDIAEHTHDQVVLFIKASCERHSGELMLLVRPNA
PTPL1	515030	1	PEREITLVNLKKDAKYGLGFQIIGGEKMGRLDLGIFISSVAPGGPADFHGCLKPGDRLISVNSVSLEGVSHHAAIEILQNAPEDVTLVISQPKEKISKVPSTPVHL
PTPL1	515030	2	GDIFEVELAKNDNSLGISVTGGVNTSVRHGGIYVKAVIPQGAAESDGRIHKGDRVLAVNGVSLEGATHKQAVETLRNTGQVVHLLLEKGQSPTSK
PTPL1	515030	3	TEENTFEVKLFKNSSGLGFSFSREDNLIPEQINASIVRVKKLFAGQPAAESGKIDVGDVILKVNGASLKGLSQQEVISALRGTAPEVFLLLCRPPPGVLPEIDT
PTPL1	515030	4	ELEVELLITLIKSEKASLGFTVTKGNQRIGCYVHDVIQDPAKSDGRLKPGDRLIKVNDTDVTNMTHTDAVNLLRAASKTVRLVIGRVLELPRIPMLPH
PTPL1	515030	5	MLPHLLPDITLTCNKEELGFSLCGGHDSLYQVVYISDINPRSVAAIEGNLQLLDVIHYVNGVSTQGMTLEEVNRALDMSLPSLVLKATRNDLPV
RGS12	3290015	1	RPSPPRVRSVEVARGRAGYGFTLSGQAPCVLSCVMRGSPADFVGLRAGDQILAVNEINVKKASHEDVVKLIGKCSGVLHMVIAEGVGRFESCS
Rhophilin-like	14279408	1	ISFSANKRWTPPRSIRFTAEEGDLGFTLRGNAPVQVHFLDPYCSASVAGAREGDYIVSIQLVDCKWLTLSEVMKLLKSFGEDEIEMKVVSLLDSTSSMHNKS
Serine Protease	2736914	1	RGEKKNSSSGISGSQRRYIGVMMLTLSPSILAELQLREPSFPDVQHGVLIHKVILGSPAHRAGLRPGDVILAIGEQMVQNAEDVYEAVRTQSQLAVQIRRGRET

Gene Name	GI	Domain Number	Sequence
Shank 1	6049185	1	EEKTVVLQKKDNEGFGFVLRGAKADTPIEEFTPTPAFPALQYLESVDEGGVAWQAGLRTGDFLIEVNNENVVKVGHRQVVNMIRQGGNHLVLKVVTVTRNLDPDDTARKKA
Shank 3	*	1	SDYVIDDKVAVLQKRDHEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAWRAGLRTGDFLIEVNGVNVVKVGHKQVVALIRQGGNRLVMKVVSV
SIP1	2047327	1	IRLCRLVRGEQGYGFHLHGEKGRRGQFIRRVEPGSPAEAAALRAGDRLVEVNGVNVEGETHHQVVQRIKAVEGQTRLLVVDQN
SIP1	2047327	2	IRHLRKGPQGYGFNLHSDKSRPGQYIRSVDPGSPAARSGLRAQDRLIEVNGQNVEGLRHAEVVASIKAREDEARLLVVDPETDE
SITAC-18	8886071	1	PGVREIHLCKDERGKTGLRLRKVDQGLFVQLVQANTPASLVGLRFGDQLLQIDGRDCAGWSSHKAHQVVKKASGDKIVVVVRDRPFQRTVTM
SITAC-18	8886071	2	PFQRTVTMHKDSMGHVGFVIKKGKIVSLVKGSSAARNGLLTNHYVCEVDGQNVIGLKDKKIMEILATAGNVVTLTIIPSVIYEHIVEFIV
SYNTENIN	2795862	1	LEIKQGIREVILCKDQDGKIGLRLKSIDNGIFVQLVQANSPASLVGLRFGDQVLQINGENCAGWSSDKAHKVLKQAFGEKITMRIHRD
SYNTENIN	2795862	2	RDRPFERTITMHKDSTGHVGFIFKNGKITSIVKDSSAARNGLLTEHNICEINGQNVIGLKDSQIADILSTS
Syntrophin 1 alpha	1145727	1	QRRRVTVRKADAGGLGISIKGGRENKMPILISKIFKGLAADQTEALFVGDAILSVNGEDLSSATHDEAVQVLKKTGKEVVLEVKYMKDVSPYFK
Syntrophin beta 2	476700	1	IRVVKQEAGGLGISIKGGRENRMPILISKIFPGLAADQSRALRLGDAILSVNGTDLRQATHDQAVQALKRAGKEVLLEVKFIREFIVTD
Syntrophin gamma 1	9507162	1	EPFYSGERTVTIRRQTVGGFGLSIKGGAEHNIPVVVSKISKEQRAELSGLLFIGDAILQINGINVRKCRHEEVVQVLRNAGEEVTLTVSFLKRAPAFLKLP
Syntrophin gamma 2	9507164	1	SHQGRNRRTVTLRRQPVGGLGLSIKGGSEHNVPVVISKIFEDQAADQTGMLFVGDAVLQVNGIHVENATHEEVVHLLRNAGDEVTITVEYLREAPAFLK
TAX2-like protein	3253116	1	RGETKEVEVTKTEDALGLTITDNGAGYAFIKRIKEGSIINRIEAVCVGDSIEAINDHSIVGCRHYEVAKMLRELPKSQPFTLRLVQPKRAF
TIAM 1	4507500	1	HSIHIEKSDTAADTYGFSLSSVEEDGIRRLYVNSVKETGLASKKGLKAGDEILEINNRAADALNSSMLKDFLSQPSLGLLVRTYPELE
TIAM 2	6912703	1	PLNVYDVQLTKTGSVCDFGFAVTAQVDERQHLSRIFISDVLPDGLAYGEGLRKGNEIMTLNGEAVSDLDLKQMEALFSEKSVGLTLIARPPDTKATL
TIP1	2613001	1	QRVEIHKLRQGENLILGFSIGGGIDQDPSQNPFSEDKTDKGIYVTRVSEGGPAEIAGLQIGDKIMQVNGWDMTMVTHDQARKRLTKRSEEVVRLLVTRQSLQ
TIP2	2613003	1	RKEVEVFKSEDALGLTITDNGAGYAFIKRIKEGSVIDHIHLISVGDMIEAINGQSLLGCRHYEVARLLKELPRGRTFTLKLTEPRK
TIP33	2613007	1	HSHPRVVELPKTDEGLGFNVMGGKEQNSPIYISRIIPGGVAERHGGLKRGDQLLSVNGVSVEGEHHEKAVELLKAAKDSVKLVVRYTPKVL
TIP43	2613011	1	ISNQKRGVKVLKQELGGLGISIKGGKENKMPILISKIFKGLAADQTQALYVGDAILSVNGADLRDATHDEAVQALKRAGKEVLLEVKYMREATPYV
X-11 beta	3005559	1	IHFSNSENCKELQLEKHKGEILGVVVVESGWGSILPTVILANMMNGGPAARSGKLSIGDQIMSINGTSLVGLPLATCQGIIKGLKNQTQVKLNIVSCPPVTTVLIK

Gene Name	GI	Domain Number	Sequence
X-11 beta	3005559	2	IPPVTTVLIKRPDLKYQLGFSVQNGIICSLMRGGIAERGGVRVGHRIIEINGQSVVATAHEKIVQALSNSVGEIHMKTMPAAMFRLLTGQENSS
ZO-1	292937	1	IWEQHTVTLHRAPGFGFGIAISGGRDNPHFQSGETSIVISDVLKGGPAEGQLQENDRVAMVNGVSMDNVEHAFAVQQLRKSGKNAKITIRRKKKVQIPNSS
ZO-1	292937	2	ISSQPAKPTKVTLVKSRKNEEYGLRLASHIFVKEISQDSLAARDGNIQEGDVVLKINGTVTENMSLTDAKTLIERSKGKLKMVVQRDRATLLNSS
ZO-1	292937	3	IRMKLVKFRKGDSVGLRLAGGNDVGIFVAGVLEDSPAAKEGLEEGDQILRVNNVDFTNIIREEAVLFLLDLPKGEEVTILAQKKKDVFSN
ZO-2	12734763	1	LIWEQYTVTLQKDSKRGFGIAVSGGRDNPHFENGETSIVISDVLPGGPADGLLQENDRVVMVNGTPMEDVLHSFAVQQLRKSGKVAAIVVKRPRKV
ZO-2	12734763	2	RVLLMKSRANEEYGLRLGSQIFVKEMTRTGLATKDGNLHEGDIILKINGTVTENMSLTDARKLIEKSRGKLQLVVLRDS
ZO-2	12734763	3	HAPNTKMVRFKKGDSVGLRLAGGNDVGIFVAGIQEGTSAEQEGLQEGDQILKVNTQDFRGLVREDAVLYLLEIPKGEMVTILAQSRADVY
ZO-3	10092690	1	IPGNSTIWEQHTATLSKDPRRGFGIAISGGRDRPGGSMVVSDVVPGGPAEGRLQTGDHIVMVNGVSMENATSAFAIQILKTCTKMANITVKRPRRIHLPAEFIVT
ZO-3	10092690	2	QDVQMKPVKSVLVKRRDSEEFGVKLGSQIFIKHITDSGLAARHRGLQEGDLILQINGVSSQNLSLNDTRRLIEKSEGKLSLLVLRDRGQFLVNIPNSS
ZO-3	10092690	3	RGYSPDTRVVRFLKGKSIGLRLAGGNDVGIFVSGVQAGSPADGQGIQEGDQILQVNDVPFQNLTREEAVQFLLGLPPGEEMELVTQRKQDIFWKMVQSEFIV

*: No GI number for this PDZ domain containing protein - it was computer cloned by J.S. using rat Shank3 seq against humen genomic clone AC000036.
In silico spliced together nt6400-6496, 6985-7109, 7211-7400 to create hypothetical human Shank3.

AVC ID	PL	Peptide Optimal Conc	PDZ	PDZ Domain	Protein Optimal Conc	Classification
AA02.1	Clasp-2	0	PSD95	1,2,3	0	2
	Clasp-2	0	NeDLG	1,2	0	2
AA10	CD46	0	Mint 1	1,2	0	1
	CD46	0	KIAA807		0	4
	CD46	0	KIAA0807(S)	1	0	5
AA13	CD95 (fas)	0	PSD95	1,2,3	0	1
	CD95 (fas)	0	NeDLG	1,2	0	1
	CD95 (fas)	0	DLG1	1,2	0	2
AA22	DNAM-1	0	PSD95	1,2,3	0	2
	DNAM-1	0	NeDLG	1,2	0	2
	DNAM-1	0	DLG1	1,2	0	1
AA29.3	IL-8RB	0	PSD95	1,2,3	0	1
	IL-8RB	0	KIAA0807(S)	1	0	1
AA216	NMDA R2C	0	PSD95	1,2,3	0	1
	NMDA R2C	0	NeDLG	1,2	0	2
	NMDA R2C	0	DLG1	1,2	0	1
AA07	CD34	0	KIAA807		0	5
	CD34	0	KIAA0807(S)	1	0	3
AA30	LPAP	0	KIAA0807(S)	1	0	5
	LPAP	0	Mint 1	1,2	0	1
	LPAP	5	TIP1	1	5	5
AA36	Neuroligin	0	KIAA0807(S)	1	0	3
AA40	Dock2	0	KIAA0807(S)	1	0	4
	Dock2	0	KIAA807		0	5
AA45	BLR-1	0	KIAA807		0	2
	BLR-1	1	KIAA0807(S)	1	0.3	2
	BLR-1	0	PDZK1	2,3,4	0	1
	BLR-1	0	KIAA0561	1	0	1
AA56	Tax	0	TIP1	1	0	5
	Tax	0	KIAA0807(S)	1	0	5
	Tax	0	KIAA807		0	5
	Tax	0	DLG1	1,2	0	5
	Tax	0	PSD95	1,2,3	0	5
	Tax	0	NeDLG	1,2	0	5
AA58	PAG	0	KIAA807		0	5
	PAG	0.35	KIAA0807(S)	1	0.5	5

WHAT IS CLAIMED IS

1. A method of modulating a biological function of a cell, comprising introducing into the cell an agent that alters binding between a PDZ protein and a PL protein in the cell, whereby the biological function is modulated in the cell, and wherein the PDZ protein and PL protein are a binding pair as specified in Table 2.

2. The method of claim 1, wherein the PDZ protein is a protein kinase, a guanalyte kinase, a tyrosine phosphatase or a serine phosphatase.

3. The method of claim 1, wherein the PDZ protein is a LIM protein or a guanine exchange factor.

4. The method of claim 1, wherein the PDZ protein is viral oncogene interacting protein.

5. The method of claim 1, wherein the PL protein is a T-cell surface receptor or a B-cell surface receptor.

6. The method of claim 1, wherein the PL protein is a natural killer cell surface receptor, a monocyte cell surface receptor, or a granulocyte cell surface receptor.

7. The method of claim 1, wherein the PL protein is an endothelial cell surface receptor.

8. The method of claim 1, wherein the PL protein is a G-protein linked receptor or a regulator of G-protein signaling.

9. The method of claim 1, wherein the PL protein is an adhesion protein or a tight junction integral membrane protein.

10. The method of claim 1, wherein the PL protein is a viral oncogene.

11. The method of claim 1, wherein the PL protein is neuron membrane transport protein.

12. The method of claim 1, wherein the PL protein is a receptor kinase.

13. The method of lcaim 1, wherein the PDZ protein is an ion channel or transporter protein.

14. The method of claim 1, wherein the PL protein is a tumor suppressor protein.

15. The method of claim 1, wherein the agent is a polypeptide comprising at least the two carboxy-terminal residues of the PL protein.

16. The method of claim 15, wherein the agent comprises at least the three carboxy-terminal residues of the PL protein.

17. The method of claim 1, wherein the agent is a small molecule or peptide mimetic of at least the two carboxy terminal residues of the PL protein.

18. The method of claim 1, wherein the agent is an antagonist that inhibits binding between the PDZ protein and PL protein binding pair.

19. The method of claim 1, wherein the agent is an agonist that promotes binding between the PDZ protein and the PL protein binding pair.

20. The method of claim 1, wherein the method is conducted in vitro.

21. A method of determining whether a test compound is a modulator of binding between a PDZ protein and a PL protein, comprising:

(a) contacting under suitable binding conditions (i) a PDZ –domain polypeptide having a sequence from the PDZ protein, and (ii) a PL peptide, wherein

the PL peptide comprises a C-terminal sequence of the PL protein,

the PDZ –domain polypeptide and the PL peptide are a binding pair as specified in Table 2; and

contacting is performed in the presence of the test compound; and

(b) detecting formation of a complex between the PDZ-domain polypeptide and the PL peptide, wherein

(i) presence of the complex at a level that is statistically significantly higher in the presence of the test compound than in the absence of test compound is an indication that the test compound is an agonist, and

(ii) presence of the complex at a level that is statistically significantly lower in the presence of the test compound than in the absence of test compound is an indication that the test compound is an antagonist.

22. The method of claim 21, wherein complex is detected in both the absence and presence of test compound.

23. A modulator of binding between a PDZ protein and a PL protein, wherein the modulator is

(a) a peptide comprising at least 3 residues of a C-terminal sequence of a PL protein, and wherein the PDZ protein and the PL protein are a binding pair as specified in Table 2; or

(b) a peptide mimetic of the peptide of section (a); or

(c) a small molecule having similar functional activity as the peptide of section (a) with respect to the PDZ and PL protein binding pair.

24. The modulator of claim 23 that is agonist.

25. The modulator of claim 23 that is an antagonist.

26. A pharmaceutical composition comprising a modulator of claim 23.

27. A method of treating a disease correlated with binding between a PDZ protein and a PL protein, the method comprising administering a therapeutically effective amount of a modulator of claim 23.

28. The method of claim 27, wherein the disease is selected from the group consisting of a neurological disease, an immune response disease, a muscular disease, and a cancer.

29. The method of claim 27, wherein the modulator is administered to a non-human animal.

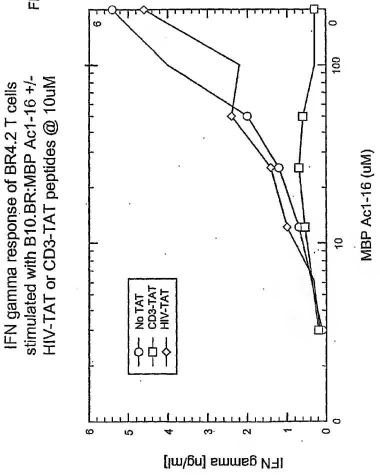
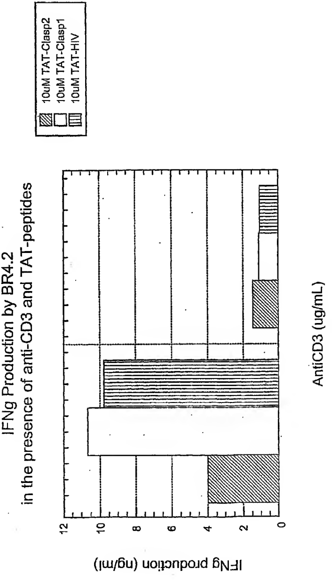

IFN gamma response of BR4.2 T cells
stimulated with B10.BR:MBP Ac1-16 +/-
HIV-TAT or CD3-TAT peptides @ 10uM
No TAT
CD3-TAT
HIV-TAT
IFN gamma [ng/ml]
0
1
2
3
4
5
6
0
10
100
MBP Ac1-16 (uM)

FIG. 1A

FIG . 1B


IFNg Production by BR4.2
in the presence of anti-CD3 and TAT-peptides
10uM TAT-Clasp2
10uM TAT-Clasp1
10uM TAT-HIV
IFNg production (ng/ml)
12
10
8
6
4
2
0
AntiCD3 (ug/mL)

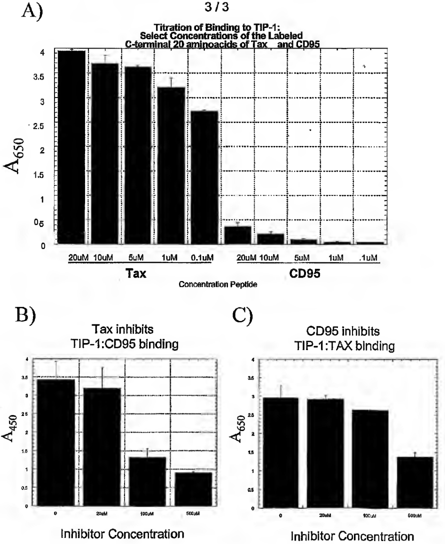

A)
Titration of Binding to TIP-1:
Select Concentrations of the Labeled
C-terminal 20 aminoacids of Tax and CD95
A_{650}
20uM 10uM 5uM 1uM 0.1uM
20uM 10uM 5uM 1uM .1uM
Tax
CD95
Concentration Peptide
B)
Tax inhibits
TIP-1:CD95 binding
A_{450}
Inhibitor Concentration
C)
CD95 inhibits
TIP-1:TAX binding
A_{650}
Inhibitor Concentration

FIG. 2